

تحليل البيانات

الدكتور حيان ديب



ISSN: 2617-989X



Books & Refrences

تحليل البيانات

د.حيان ديب

من منشورات الجامعة الافتراضية السورية

الجمهورية العربية السورية 2018

Data Analysis

Hayan Deeb

Publications of the Syrian Virtual University (SVU)

Syrian Arab Republic, 2018

ISSN: 2617-989X

الفهرس

الفصل الأول الاستبانة والمقاييس

Chapter 1: Questionnaire and scales

الكلمات المفتاحية:

الاستبانة، القياس، المقياس الاسمي، المقياس الترتيبي، المقياس المدرج، مقياس النسب.

ملخص:

يشرح الفصل مفهوم القياس وأنواع المقاييس الأساسية ابتداء بالمقاييس الاسمية مروراً بالمقاييس الترتيبية والمدرجة وانتهاء بمقاييس النسب. كما يميز الفصل بين مقاييس المقارنة ومقاييس اللامقارنة. يتناول الفصل أيضاً بعض الأمور الأساسية المتعلقةبتصميم الاستبانة وكيفية التعامل مع عدم قدرة المجيب أو عدم رغبته في الإجابة وترتيب الأسئلة والاختبار الأولي للاستبانة.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- التعرف على أنواع المقاييس الأساسية
- فهم تقنيات القياس القائمة على المقارنة أو على اللامقارنة
- إدراك كيفية تصميم الاستبانة
- معرفة كيفية التعامل مع عدم قدرة المجيب على الإجابة أو عدم رغبته في الإجابة
- فهم كيفية اختبار الاستبانة

الفصل الأول: الاستبانة والمقاييس

تعتبر عملية **القياس** إحدى المراحل الأساسية في البحوث التي تستخدم الأساليب الإحصائية للخروج بالنتائج والقرارات. كما أنها تعتبر الأساس الذي يحكم درجة الثقة والصدق فيما يمكن أن يصل إليه الباحث من نتائج وخلاصات.

وكثيراً ما يتم استخدام المقاييس ضمن الاستبانات التي يتم تصميمها وتوزيعها للحصول على البيانات التي سيتم معالجتها إحصائياً في مرحلة لاحقة.

سنتناول في هذا الفصل بعض الأمور الأساسية المتعلقة بتصميم الاستبانة كما سنتعرف على أنواع المقاييس الأكثر انتشاراً واستخداماً من قبل الباحثين ضمن الاستبانات.

1. أنواع المقاييس *Types of scales*

يعرف القياس مفاهيمياً بأنه الإجراء الذي يتم بواسطته تحديد قيم معيارية (أرقام أو حروف أو رموز) لحالات التغيّر التي يأخذها المتغير محل القياس والخصائص التي تتصف بها كل حالة. وتصنف المقاييس بحسب مستوى المعلومة التي توفرها ضمن أربع مجموعات هي:

- المقاييس الاسمية
- المقاييس الترتيبية
- المقاييس المدرجة
- مقاييس النسب

1-1. المقاييس الاسمية Nominal scales

تكون القيم أو الأرقام التي يتكون منها المقياس الاسمي *Nominal Scale* مجرد علامات مميزة أو إشارات للتعريف بمفردات عينة الدراسة أو الأشياء أو لتصنيفها.

عندما يتم استخدام المقياس الاسمي للتعريف *Identification* يكون هناك توافق تام بين الأرقام والأشياء أو الأفراد. وكمثال على هذا المقياس نذكر:

- الأرقام الوطنية على البطاقات الشخصية

- الأرقام التي يتم إعطاؤها للاعبي كرة القدم
- أرقام الطلاب التي يتم إعطاؤها لهم عند التسجيل
- أرقام الحسابات في البنوك
- الأرقام التي يتم إعطاؤها لتمييز المجيبين.

يكون لدينا في النهاية مجموعة من الأرقام التي يشير كل منها إلى شيء/فرد. تساعد هذه المطابقة بين الأرقام والأشياء/الأفراد على وصف كل فرد/شيء على حده. ولا بد أن نتأكد من عدم وجود فردين/شيئين يحملان نفس الرقم.

أما عندما يتم استخدام المقياس الاسمي بهدف التصنيف *Classification* فإن الأرقام المختارة تستخدم كعناوين لتمييز المجموعات أو الفئات كالجنس أو المستوى التعليمي أو الحالة الاجتماعية،....

وسواء استخدمت الأرقام في المقاييس الاسمية للتعريف أو التصنيف فإنها لا تمتلك صفات الأرقام في المعاملات الرياضية والحسابية. فهي لا تعكس أي الأشياء أكبر أو أي الخصائص أفضل. فاللاعب الذي يحمل الرقم واحد مثلاً لا يحكم عليه من خلال رقمه على أنه أقل شأناً من اللاعب الذي يحمل الرقم اثنين. كما يمكن إعطاء الرقم 1 للذكور و2 للإناث أو العكس من دون أن تعني الأرقام شيئاً عدا كونها تمثل عناوين للمجموعات والأفراد. أي أن الرقمين لا يعطيان المعنى الحقيقي للأرقام وبالتالي لا يمكن إجراء العمليات الحسابية على هذه الأرقام.

من خلال خصائص المقاييس الاسمية نستنتج أن العمليات الوحيدة التي تسمح بها الأرقام تقوم على العد *counting* أو التكرار *frequency*. فلا يمكن مثلاً حساب الوسط الحسابي للجنس ولكن يمكن حساب عدد الذكور وعدد الإناث في العينة.

تعتبر المقاييس الاسمية من أقل أنواع المقاييس من حيث الغنى بالمعلومات. وهي تسمح بعدد محدود من العمليات الإحصائية عند التحليل كحساب النسب المئوية والتكرار والمنوال واختبار كاي مربع.

1-2. المقاييس الترتيبية Ordinal scales

كما يوحي اسمها تتضمن المقاييس الترتيبية *Ordinal scales* بعداً ترتيبياً. ويتطلب استخدامها قدرة على التمييز بين مفردات عينة الدراسة طبقاً لخاصية معينة. فمثلاً يستطيع الباحث ترتيب مجموعة من أصناف مساحيق الغسيل طبقاً لقدرتها على التنظيف. فإذا أعطى رقم 1 لأكثر الأصناف قدرة على التنظيف، ورقم 2 للصنف الأقل، وهكذا... فإنه سينتج لديه في النهاية مقياس ترتيبي لمساحيق الغسيل. كما يمكن

استخدام هذا النوع من المقاييس لتصنيف العاملين بناء على مهاراتهم القيادية من الأكثر إلى الأقل.

يسمح هذا النوع من المقاييس بالترتيب ولكنه لا يوضح الفروقات أو المسافات بين الأمور التي يتم ترتيبها. بمعنى أن المقياس لا يساعد في توضيح الفرق بين الأشياء/الأفراد طبقاً للخاصية التي يتم الترتيب على أساسها حيث نستطيع أن نقول إن *A* أكبر من *B* ولكننا لا نستطيع أن نعرف المسافة بين *A* و*B* باستخدام هذه المقاييس. لا تعني الفروقات بين الأرقام إذاً أكثر من مجرد اختلاف في الترتيب أو الأهمية.

تشبه المقاييس الترتيبية المقاييس الاسمية من حيث البيانات التي توفرها للباحث ولكنها تعتبر أكثر غنى بالمعلومات. يمكن استخدام طرق الإحصاء الوصفي المتعلقة بالتكرار وغيرها من طرق التلخيص الإحصائي المتعلقة بترتيب المفردات محل البحث مثل: الوسيط والنسب المئوية والمنوال والربيعيات. أما الوسط الحسابي العادي للعمليات فليس له أي دلالة تفسيرية بالنسبة للبيانات التي يوفرها المقياس الترتيبي.

1-3. المقاييس المدرجة أو الفئوية Interval scales

تنطوي عملية القياس باستخدام المقاييس المدرجة *Interval Scales* على وجود وحدة قياس ثابتة في المقياس. يساعد المقياس على صياغة عبارات ذات دلالات معينة حول الفروقات الموجودة بين الأفراد.

يقدم هذا المقياس كل المعلومات التي يقدمها المقياس الترتيبي بالإضافة إلى تقديمه فرصة مقارنة وتقييم الفروقات بين الأشياء. وتكون الفروقات أو المسافات متساوية بين درجات المقياس. كما أننقطة الصفر على المقياس تمثل قيمة تخمينية أو اعتبارية فالصفر في هذا النوع من المقاييس لا يعني عدم توافر الصفة. وتعتبر موازين الحرارة أمثلة للمقاييس المدرجة فدرجة الحرارة صفر لا تعني عدم وجود حرارة.

تضم الطرق الإحصائية التي يمكن استخدامها مع المقاييس المدرجة كل الطرق التي يمكن استخدامها مع المقاييس الاسمية والترتيبية بالإضافة إلى إمكانية حساب الوسط الحسابي والانحراف المعياري والارتباط الخطي ومجموعة واسعة أخرى من الطرق الإحصائية.

1-4. مقاييس النسب Ratio scales

تعتبر مقاييس النسب *Ratio Scale* ذات فاعلية عالية لأغراض كثيرة في الأبحاث التطبيقية. ويعني

الصفر في هذه المقاييس عدم وجود الصفة (المعنى الحقيقي للصفر) فالزمن صفر يعني عدم وجود زمن وكذلك المسافة صفر. وتعني النسب المتساوية بين قيم المقياس نسباً متساوية بين مفردات عينة البحث. ومن أمثلة مقاييس النسب نذكر الطول والوزن والتكاليف والحصة السوقية والمبيعات وغيرها.

ويتم التعامل عادة مع المقاييس المدرجة/الفئوية ومقاييس النسب بالطرق نفسها ويطلق عليها عادة اسم المقاييس **الكمية أو المستمرة أو القياسية** بحسب المرجع المستخدم.

المقياس				
اسمي	أرقام قمصان العدائين	3	8	7
ترتيبي	ترتيب وصول اللاعبين	First place	Second place	Third place
مدرج	تقييم الأداء على مقياس من ١٠ درجات	9.6	9.1	8.2
نسب	الزمن المستغرق لإنهاء السباق	13.4	14.1	15.2

الشكل 1-1. أنواع المقاييس

2. تقنيات القياس Scaling technics

يمكن تمييز عائلتين من المقاييس بناء على تقنيات القياس المستخدمة وهي مقاييس المقارنة ومقاييس اللامقارنة.

2-1. مقاييس المقارنة Comparative scales

تقوم تقنية القياس في مقاييس المقارنة *Comparative scales* على المقارنة المباشرة بين الأشياء المبحوثة أو المختبرة. وتتمتع هذه المقاييس بصفات المقاييس الاسمية والترتيبية فقط. ومن أشكالها نذكر على سبيل المثال لا الحصر مقاييس المقارنة الثنائية ومقاييس المقارنة الترتيبية.

2-1-1. مقاييس المقارنة الثنائية Paired comparison scales

يتم في مقاييس المقارنة الثنائية *Paired Comparison Scaling* عرض شيئين على المجيب ويطلب منه اختيار شيء منهما. تكون البيانات الناتجة ذات طبيعة ترتيبية *ordinal*. وينصح ألا يكون عدد الثنائيات التي نرغب بمقارنتها كبيراً.

مثال: سنعرض عليكم عدة ثنائيات من العلامات التجارية للشامبو. في كل ثنائية، نرجو منكم اختيار العلامة التجارية التي تفضلونها للاستخدام الشخصي.

	Jhirmack	*Finesse*	*Vidal Sassoon*	*Head & Shoulders*	*Pert*
Jhirmack		0	0	1	0
Finesse	1[a]		0	1	0
Vidal Sassoon	1	1		1	1
Head & Shoulders	0	0	0		0
Pert	1	1	0	1	
عدد التفضيلات	3	2	0	4	1

a. يعني 1 أن العلامة التجارية في العمود تم تفضيلها على العلامة الظاهرة في السطر.
يعني 0 أن العلامة التجارية في السطر تم تفضيلها على العلامة الظاهرة في العمود
b. عدد التفضيلات = مجموع قيم العمود

2-1-2. مقاييس الترتيب Rank order scales

يتم في مقاييس الترتيب *Rank order scaling* عرض عدة أشياء دفعة واحدة على المجيب ويطلب منه ترتيبها بناء على معيار محدد. يعتبر هذا القياس أقرب إلى الحياة العملية من القياس باستخدام مقياس المقارنة الثنائي.

مثال: رتب العلامات التجارية التالية لمعاجين الأسنان بحسب تفضيلها. ابدأ بإعطاء القيمة 1 للعلامة الأكثر تفضيلاً بالنسبة لك ثم 2 للعلامة التي تليها من حيث التفضيل واستمر بهذه الطريقة وصولاً إلى إعطاء القيمة 10 للعلامة الأقل تفضيلاً بالنسبة لك.

العلامة التجارية	الترتيب
1. Crest	______
2. Colgate	______
3. Aim	______
4. Gleem	______
5. Sensodyne	______
6. Ultra Brite	______
7. Close Up	______
8. Pepsodent	______
9. Plus White	______
10. Stripe	______

2-2. مقاييس اللامقارنة Non-comparative scales

يتم في مقاييس اللامقارنة قياس خصائص الشيء المبحوث بمعزل عن الأشياء الأخرى. وتمتلك البيانات الناتجة عن هذه المقاييس خصائص بيانات المقاييس المدرجة والنسب. ونذكر من أنواعها: المقاييس المستمرة/التخطيطية ومقياس ليكرت ومقياس تباين المعاني.

2-2-1. المقياس المستمر Continuous rating scale

يعرف المقياس المستمر *continuous rating scale* بالمقياس التخطيطي *graphic rating scale* أيضاً. تتم الإجابة هنا من خلال وضع علامة على الخط المستمر/المستقيم الدال على المقياس. ويمكن أن يكون الخط أفقياً أو عمودياً. وقد أصبح هذا المقياس أكثر شعبية مع تطور الوسائل التكنولوجية المستخدمة في الاستبيان (الكمبيوتر والانترنت).

مثال:

كيف تقيم الجامعة أداء الجامعة الافتراضية؟

ممتاز ----------------------------I------- ضعيف

ممتاز ----------------------------I------- ضعيف

0 10 20 30 40 50 60 70 80 90 100

2-2-2. مقياس ليكرت Likert scale

سمي مقياس ليكرت *Likert scale* بهذا الاسم نسبة إلى مصممه *Likert Rensis*. يقوم هذا المقياس على إعطاء درجة الموافقة أو عدم الموافقة على مجموعة من العبارات المتعلقة بالشيء المدروس.

- 1 = غير موافق بشدة
- 2 = غير موافق
- 3 = محايد
- 4 = موافق
- 5 = موافق بشدة

مثال: أعط درجة موفقتك على كل من العبارات التالية:

	غير موافق بشدة	**غير موافق**	**محايد**	**موافق**	**موافق بشدة**
1. تنتج *Renault* سيارات عالية الجودة	**1**	**2**	**3**	**4**	**5**
2. خدمة ما بعد البيع سيئة لدى *Renault*	**1**	**2**	**3**	**4**	**5**
3. لا أحب إعلانات سيارات *Renault*	**1**	**2**	**3**	**4**	**5**

3-2-2. مقياس تباين المعاني Semantic differential scale

يتكون مقياس تباين المعاني *Semantic Differential Scale* من 7 درجات (عادةً) حيث يتضمن في الطرف الأول منه صفة مناقضة تماماً للصفة الموجودة في الطرف الآخر.

مثال:

أعتقد أن رئيسي في العمل:

: ضعيف *:--:--:-X-:--:--:--:--* قوي

: لا يعتمد عليه *:--:-X-:--:--:--:--:--* يعتمد عليه

: فات عصره *:-X-:--:--:--:--:--:--* عصري

3. تصميم الاستبانة Questionnaire Design

تعرف الاستبانة على أنها مجموعة من الأسئلة المصممة لجمع أكبر قدر من البيانات اللازمة عن المشكلة أو الظاهرة محل البحث لتحقيق الأهداف الأساسية للبحث. وفي الحقيقة فإنه لا توجد ما تسمى بالاستبانة النمطية أو المثالية والتي يمكن أن تكون صالحة لكافة الأوضاع التي تجري فيها البحوث. لكن ونظراً لأن تصميم صحيفة الاستبانةيعتبر محدداً هاماً لدقة البيانات فإن الباحث يجب أن يولي إعدادها وتصميمها الاهتمام الكافي الذي يرقى بها إلى مستوى الدقة المطلوب.

3-1. تحديد محتوى الأسئلة Question content

يتوجب عند تحديد محتوى الأسئلة التفكير في الحاجة الفعلية للسؤال والابتعاد عن السؤال عن أكثر من شيء ضمن السؤال الواحد وتجنب الأسئلة الإيحائية.

3-1-1. الحاجة للسؤال Need for the question

إذا لم يكن هناك حاجة للبيانات التي ستنتج عن السؤال، ينصح بحذف السؤال. إلا أنه يمكن في بعض الأحيان أن تطرح أسئلة حيادية لشد انتباه المجيب وزيادة إمكانية مشاركته في الاستبيان وخصوصاً عندما يكون الموضوع حساساً أو جدلياً. كما يمكن أيضاً طرح بعض الأسئلة في البداية لإخفاء الهدف الحقيقي من البحث أو لمن ستعود النتائج. فمثلاً يمكن لاستبيان عن حواسب *IBM* أن يتضمن أسئلة عن *Dell*.

3-1-2. الحاجة لسؤال أو أكثر One or several questions

من الأخطاء الشائعة عند تصميم أسئلة الاستبانة أن يدور السؤال حول شيئين أو أكثر كما في الأمثلة التالية:

- هل تزايدت مبيعاتك من المنتج س أم تناقصت خلال الشهر الماضي؟

نعم _____ لا_____ لم تتغير ______

- هل تشتري صحيفة س لأنها سياسية واقتصادية ورياضية؟

نعم ______ لا _______

- هل تعتقد أن مشروب "كوكا كولا" لذيذ ومنعش؟

يعتبر السؤال السابق خاطئاً لدمجه محاولة الحصول على معلومات متعلقة بشيئين (1. لذيذ، 2. منعش) في سؤال واحد (يجب أن يغطي السؤال نقطة أو شيئاً واحداً فقط). للحصول على المعلومة الصحيحة يجب تحويل السؤال السابق إلى سؤالين:

- هل تعتقد أن مشروب كوكا كولا لذيذ؟
- هل تعتقد أن مشروب كوكا كولا منعش؟

3-1-3. الأسئلة الإيحائية Leading questions

يجب تجنب الأسئلة الإيحائية. وهي تلك الأسئلة التي توحي بإجابة معينة للسؤال.

مثال:

- ألا تعتقد أن تنظيف المنزل يحتاج إلى مجهود كبير؟

نعم _____ لا ______

3-2. تجاوز عدم القدرة على الإجابة Overcoming inability to answer

قد تنتج عدم القدرة على الإجابة من عدم امتلاك الفرد للمعلومة المطلوبة أو من عدم قدرته على التذكر.

3-2-1. امتلاك المعرفة الكافية Having the information

قد يتم سؤال الأفراد حول أمور لا يمتلكون المعلومات أو المعرفة الكافية حولها. فقد لا يمتلك الزوج معلومات عن مصروف العائلة على المواد الاستهلاكية الأساسية إذا كانت الزوجة هي من يقوم بالشراء.

من ناحية أخرى يجب أن يخلو السؤال من الكلمات الغامضة أو الألفاظ غير المعتادة للمستقصى منهم. فمثلاً تعتبر كلمة "سلعة رأسمالية" غير معتادة لعامة الناس ويختلف معناها من شخص لآخر.

تكمن المشكلة هنا في أن الأبحاث قد أظهرت أن الأفراد يميلون للإجابة على الأسئلة الواردة في الاستبيان حتى لو لم يمتلكوا المعلومات الكافية للإجابة

وقد تفيد **الأسئلة الكاشفة** ***filter questions*** في استبعاد الأفراد الذين لا يمتلكون المعلومات أو

المعرفة الكافية للإجابة على أسئلة الاستبيان كالأسئلة حول معرفة العلامة التجارية أو الخبرة في استخدام المنتج أو امتلاك معلومات عن المنتج....

وقد يفيد خيار "لا أعرف" *I don't know* في حل هذه المشكلة دون التخلي عن إجابات المجيب الأخرى الصحيحة. ويمكن إدراج هذا الخيار في حالة الشك في نقص المعلومات الكافية لدى المجيب للإجابة على بعض أسئلة الاستبيان.

3-2-2. القدرة على التذكر Ability to remember

نعتقد أحياناً بأن هناك معلومات بسيطة يستطيع غالبية الأفراد تذكرها. لكن العديد من الدراسات أظهرت أن الحقيقة قد تكون عكس ذلك.

- كم عبوة مشروب غازي استهلكتم خلال فترة معينة؟...

يفضل تحويل السؤال في هذه الحالة إلى الشكل التالي:

- وسطياً، كم مرة تستهلكون المشروبات الغازية خلال الأسبوع؟

1. _____ أقل من مرة أسبوعياً
2. _____ 1 – 3 مرات أسبوعياً
3. _____ 4 – 6 مرات أسبوعياً
4. _____ 7مرات أو أكثر أسبوعياً

كما أنه من المفيد أحياناً مساعدة المجيب على التذكر والإجابة. فقد يصعب على المجيب تذكر الإجابة على السؤالين التاليين:

- ما هي العلامة التجارية لإطارات سيارتك؟
- ما هي العلامات التجارية التي شاهدت إعلاناتها مساء البارحة؟

يمكن هنا مساعدة المجيب من خلال إعداد قائمة من أسماء العلامات التجارية ومن ثم الطلب من المجيب الإطلاع على القائمة والإجابة على السؤال التالي:

- ضع إشارة على العلامة التجارية التي شاهدت إعلاناً لها مساء البارحة.

3-3. عدم الرغبة في الإجابة Unwillingness to answer

قد يؤدي إجهاد المجيب إلى ازدياد عدم رغبته في الإجابة على أسئلة الاستبانة كما قد تقود الأسئلة الحساسة إلى تجنب الإجابة عليها من قبل المجيب.

1-3-3. الجهد المبذول للإجابة Required effort

✓ ينصح بالسعي إلى تخفيف الجهد الذي سيبذله المجيب للإجابة على السؤال.

مثال:

ما هي الأجنحة التي تقوم بالتسوق منها عادة عند زيارتك لمراكز التسوق؟

يشكل السؤال السابق مثالاً على سؤال يتطلب من المجيب بذل جهد كبير للإجابة عليه. لتقليص الجهد المبذول، يمكن طرح السؤال كمايلي:

من القائمة التالية، ضع إشارة إلى جانب الأجنحة التي تقوم بالتسوق منها عادة عند زيارتك لمراكز التسوق.

1. الألبسة النسائية ______
2. الألبسة الرجالية ______
3. ألبسة الأطفال ______
4. مستحضرات التجميل ______

.

.

9. المجوهرات ______
10. أخرى (يرجى تحديدها) ______

2-3-3. المعلومات الحساسة Sensitive information

تعتبر المعلومات الحساسة معلومات قد يرغب المجيب بإخفائها أو عدم الإجابة عليها بشكل صحيح لأنها قد تؤدي إلى حكم الآخرين عليه بشكل سلبي أو إلى التأثير على تقدير الذات بشكل سلبي. ومن أمثلة المعلومات الحساسة نذكر: المال/الدخل والحياة العائلة والمعتقدات الدينية والسياسية والتورط في أمور غير قانونية أو غير أخلاقية.... وغيرها.

وللتخفيف من هذه الظاهرة، ينصح بمايلي:

- وضع الأسئلة الحساسة في نهاية الاستبيان. قد تؤدي هذه الطريقة إلى إزالة عدم الثقة مع التقدم في الاستبيان، وخلق الرغبة في إنهاء الاستبيان من قبل المجيب.
- التمهيد للسؤال من خلال عبارة تشير إلى شيوع السلوك الذي ستسأل عنه. مثال: قبل السؤال عن القروض المصرفية يمكن عرض عبارة "تشير الدراسات الحديثة إلى معظم السوريين يقترضون من البنوك".
- استخدام أسلوب الشخص الثالث: طرح السؤال بصيغة تشير إلى إلى أنه يتعلق بأشخاص آخرين (لا يتعلق بالمجيب ولا بالسائل).
- إخفاء السؤال بين مجموعة من الأسئلة التي لا يمانع المجيب بالإجابة عليها.
- تحديد مجالات للإجابة عوضاً عن طلب الإجابة الدقيقة. فمثلاً عوضاً عن السؤال "ما هو دخلك السنوي" حدد مجالات للدخل واطلب من المجيب أن يختار المجال الذي يتوافق مع دخله.

3-4. ترتيب الأسئلة Order of questions

يمكن أن تكون الأسئلة الافتتاحية أساسية في ترغيب المجيب في متابعة الاستبيان لذا ينصح بأن تكون الأسئلة الافتتاحية بسيطة، ممتعة/ وألا تحتوي أسئلة حساسة أو مزعجة. ويمكن في بعض الأحيان البدء بأسئلة لتقييم مدى مطابقة المجيب للخصائص الواجب أن تتوفر فيه (مدخن، متزوج...)

وبشكل عام يصنف الخبراء المعلومات المضمنة في الاستبيان في ثلاث مجموعات:

- **معلومات أساسية:** معلومات مرتبطة ارتباطاً مباشراً بمشكلة البحث
- **معلومات تصنيفية:** معلومات ديموغرافية ومالية لتصنيف المجيبين وفهم النتائج
- **معلومات تعريفية:** تتضمن الاسم، العنوان، البريد الإلكتروني، رقم الهاتف.

ينصح بشكل عام بالبدء بالمعلومات الأساسية ثم المعلومات التصنيفية والانتهاء بالمعلومات التعريفية. كما ينصح بالبدء بالأسئلة العامة والانتقال نحو الأسئلة الخاصة. ويفضل وضع الأسئلة الصعبة أو الأسئلة الحساسة في مكان متأخر في الاستبيان. كما ينصح بوضع الدخل في نهاية الأسئلة التصنيفية ووضع رقم التلفون في نهاية الأسئلة التعريفية.

3-5. الاختبار الأولي للاستبانة Pretesting the questionnaire

يفضل أن يتم اختبار الاستبانة على عينة صغيرة من الأفراد الذين تتوافق خصائصهم مع خصائص العينة الرئيسية. وينصح أن يتم الاختبار من خلال المقابلة الشخصية حيث يمكن للباحث ملاحظة ردود فعل المجيب.

وتوجد طريقتان شهيرتان للاختبار هما: طريقة البروتوكول وطريقة استخلاص المعلومات.

في الطريقة الأولى "**اختبار البروتوكول** ***protocol analysis***" يطلب من المجيب أن يفكر بصوت عالٍ أثناء الإجابة على الاستبيان. ويتم عادة تسجيل ملاحظات المجيب وتحليلها لتحديد ردود فعله على مختلف أجزاء الاستبيان.

أما في الطريقة الثانية "**استخلاص المعلومات** ***debriefing***" فيبدأ الفرد بالإجابة على الاستبيان ولدى انتهائه من الإجابة يتم إعلامه أن الاستبيان الذي قام بالإجابة عليه في طور الاختبار ويشرح له معنى الاختبار. يطلب من الفرد بعد ذلك أن يشرح معنى كل سؤال وأن يشرح إجاباته وأن يتحدث عن أية مشكلات واجهته أثناء الإجابة.

أخيراً ينصح أن يعاد الاختبار بعد كل تعديل يتم على الاستبيان ويتم ذلك على عينة أخرى من الأفراد حيث يعاد الاختبار حتى زوال الحاجة إلى التعديل.

قائمة المراجع:

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام *SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Blumberg B., Cooper D.R., & Schindler P.S. (2005),* ***Business Research Methods****, Mcgraw-Hill, Berkshire.*

- *Hair J.F. JR., Bush R.P., &Ortinau D.J. (2003),* ***Marketing Research Within a Changing Information Environment****, McGraw-Hill/Irwin, 2nd Edition, USA.*

- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*

- *Malhotra N.K. &Briks D.F. (2007),* ***Marketing Research: An Applied Approach****, 3rd European Edition, Pearson Education Limited, Italy.*

- *Mooi E. &Sarstedt M. (2011),* ***A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics****, Springer, Germany.*

- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows****, 3rd Edition, McGraw Hill, USA.*

- *Zikmund W.G. &Babin B.J. (2010),* ***Essentials of Marketing Research****, 4th Edition, South-Western Cengage Learning, USA.*

مقترحات وتمارين للفصل الأول

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1. **لماذا تستخدم المقاييس الاسمية؟**

   (الحل في الفقرة: 1.1.)

2. **ما الذي يميز المقاييس الترتيبية عن المقاييس الاسمية؟**

   (الحل في الفقرة: 2.1.)

3. **ما هي المقاييس الكمية أو القياسية؟**

   (الحل في الفقرة: 4.1.)

4. **يمكن تمييز عائلتين من المقاييس بناء على تقنيات القياس المستخدمة، ما هما؟**

   (الحل في الفقرة: 2.)

5. **سنعرض عليكم عدة ثنائيات من العلامات التجارية. في كل ثنائية، نرجو منكم اختيار العلامة التجارية التي تفضلونها للاستخدام الشخصي.**

   ما نوع المقياس المستخدم في هذا السؤال؟

   (الحل في الفقرة: 1.1.2.)

6. **في أي نوع من المقاييس يتم عرض عدة أشياء دفعة واحدة على المجيب ويطلب منه ترتيبها بناء على معيار محدد؟**

   (الحل في الفقرة: 2.1.2.)

7. **ما هو نوع البيانات الناتجة عن مقاييس اللامقارنة؟**

   (الحل في الفقرة: 2.2.)

8. **ما اسم المقياس الذي يقوم على إعطاء درجة الموافقة أو عدم الموافقة على مجموعة من العبارات المتعلقة بالشيء المدروس؟**

   (الحل في الفقرة: 2.2.2.)

9. **أعتقد أن الجامعة الافتراضية:**

   سيئة :--:--:--:-X-:--:--:--:--جيدة

ما اسم المقياس المستخدم في السؤال السابق؟

(الحل في الفقرة 3.2.2.)

10. **كيف يمكن مساعدة المجيب على تذكر الإجابة على أسئلة الاستبانة؟**

(الحل في الفقرة 2.2.3.)

الفصل الثاني مدخل إلى SPSS

Chapter 2: Introduction to SPSS

الكلمات المفتاحية:

SPSS، نوافذ *SPSS*، ملفات *SPSS*، قوائم *SPSS*، إدخال البيانات في *SPSS*.

ملخص:

يقدم الفصل تعريفاً سريعاً ببيئة النظام الإحصائي *SPSS* من خلال استعراض النوافذ الأساسية المتاحة في هذا النظام وقوائمه وملفاته. كما يشرح الفصل كيفية إدخال البيانات وتعريف المتغيرات ضمن *SPSS*.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- التعرف على النوافذ الأساسية في *SPSS*.
- التعرف على قوائم *SPSS*.
- تمييز ملفات *SPSS*.
- استيعاب كيفية إدخال وترميز البيانات وتعريف المتغيرات ضمن *SPSS*.

الفصل الثاني: مدخل إلى SPSS

يقوم الكثير من المهتمين في ميادين العلوم الاقتصادية والإدارية، والتربوية والاجتماعية وغيرها بإجراء التحليلات الإحصائية لبياناتهم المختلفة، بهدف إيجاد مقاييس النزعة المركزية مثل الوسط الحسابي لمجموعة من البيانات، وحساب مقاييس التشتت وحساب معاملات الانحدار والارتباط ...إلخ. وقد سهل التطور الكبير في عالم البرمجيات الإحصائية من عمل هؤلاء المهتمين والباحثين من خلال طرح العديد من الأنظمة والبرمجيات الإحصائية وعلى رأسها الحزمة الإحصائية للعلوم الاجتماعية *Statistical Package for Social Sciences (SPSS)*.

يقدم هذا الفصل تعريفاً سريعاً ببيئة النظام الإحصائي *SPSS* حيث يشرح النوافذ والقوائم الأساسية للنظام. يقدم الفصل أيضاً معلومات أساسية تساعد على البدء باستخدام *SPSS*.

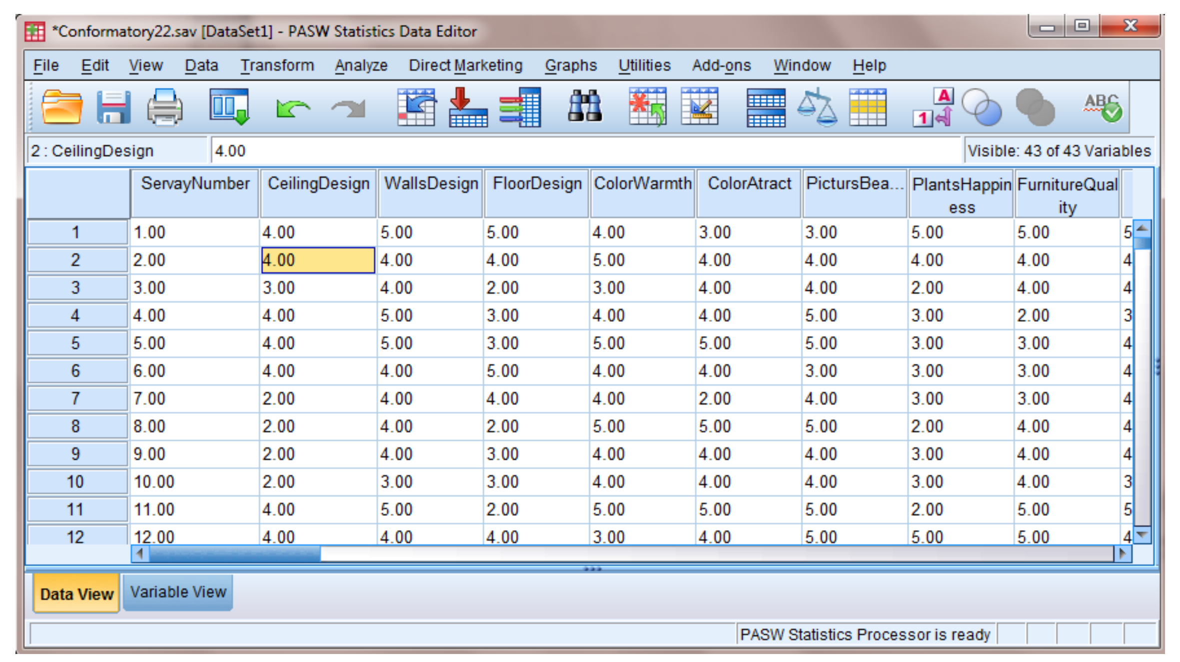
1. النوافذ المتوفرة في SPSS: *Windows available in SPSS*

يحتوي نظام *SPSS* على ثلاثة أنواع رئيسية من النوافذ هي: نافذة محرر البيانات *Data Editor Window* ونافذة المخرجات *Output Viewer* ونافذة محرر التعليمات *Syntax Editor*.

1-1. نافذة محرر البيانات Data Editor

تكون نافذة محرر البيانات *Data Editor* نشطة بشكل تلقائي عادة لدى تشغيل برنامج *SPSS* أي أنها تظهر عند فتح جلسة *SPSS* وتعرض مكونات ملف البيانات. وتستخدم هذه النافذة لتعريف وإدخال وتحرير وعرض البيانات المتعلقة بالبحث والمراد تحليلها. ويمكن من خلال هذه النافذة أن نقوم بخلق ملفات بيانات جديدة أو تعديل ملفات بيانات موجودة.

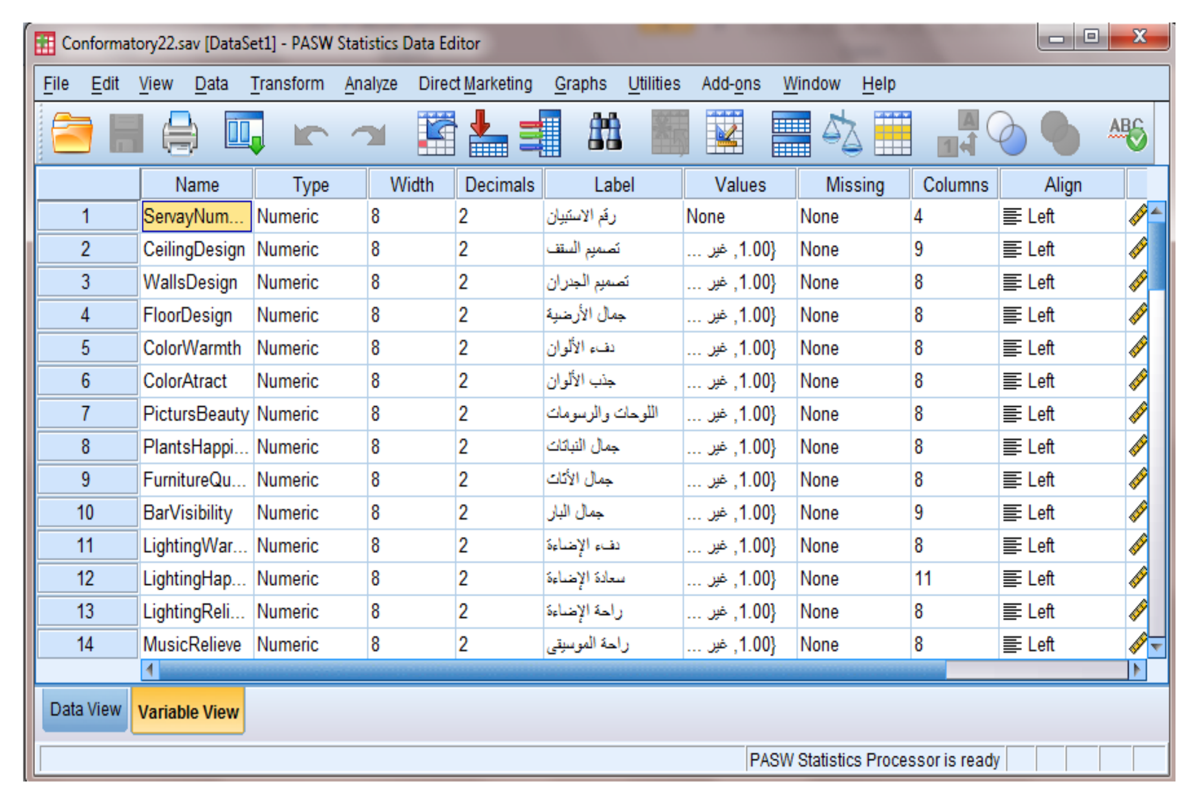
يمكن ملاحظة شريطين أسفل نافذة محرر البيانات *Data Editor* وهما شريطعرض البيانات *Data View*وشريط عرض المتغيرات *Variable View*.ويمكن التنقل أو تفعيل أي من هذين الشريطين بالنقر على اسم الشريط المراد تفعيله.



الشكل 2.1 نافذة عرض البيانات

عندما يكون شريط عرض البيانات *Data View* نشطاً فإنه يظهر البيانات المراد تحليلها والتي تم إدخالها ضمن البرنامج. وتشبه نافذة عرض البيانات إلى حد كبير شريحة *Excel*، حيث تقسم الشريحة إلى أسطر (صفوف) *cases* وأعمدة *columns*. تتضمن الصفوف (الأسطر) الحالات التي تم إجراء القياس عليها (المجيبون في حالة الاستبانة الموزعة مثلاً). أما الأعمدة فتتضمن متغيرات الدراسة حيث يحتوي كل عمود على متغير محدد.

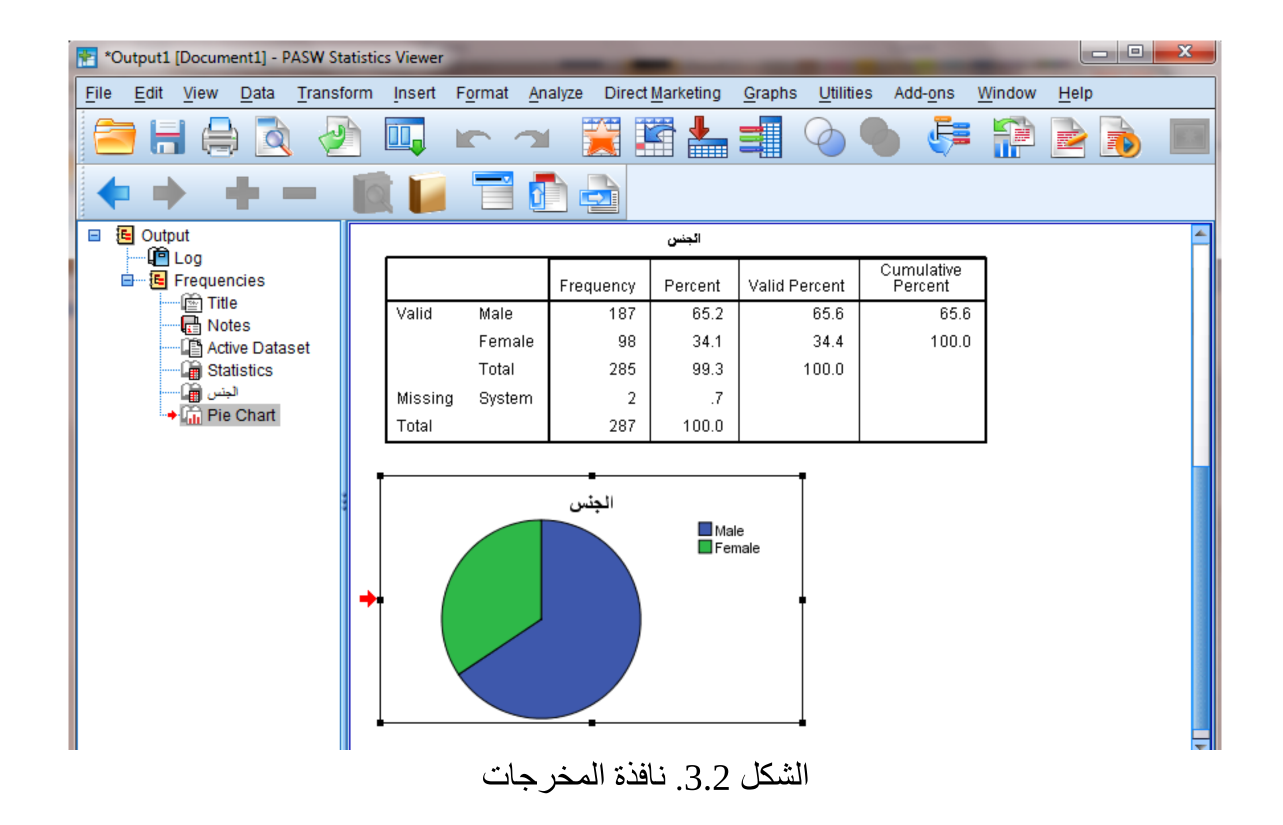
وعند النقر على شريط عرض المتغيرات *Variable View* تظهر نافذة تعريف المتغيرات التي تستخدم لعرض وتعريف المتغيرات.



شكل 2.2. نافذة عرض المتغيرات

2-1. نافذة المخرجات ***Output Viewer***

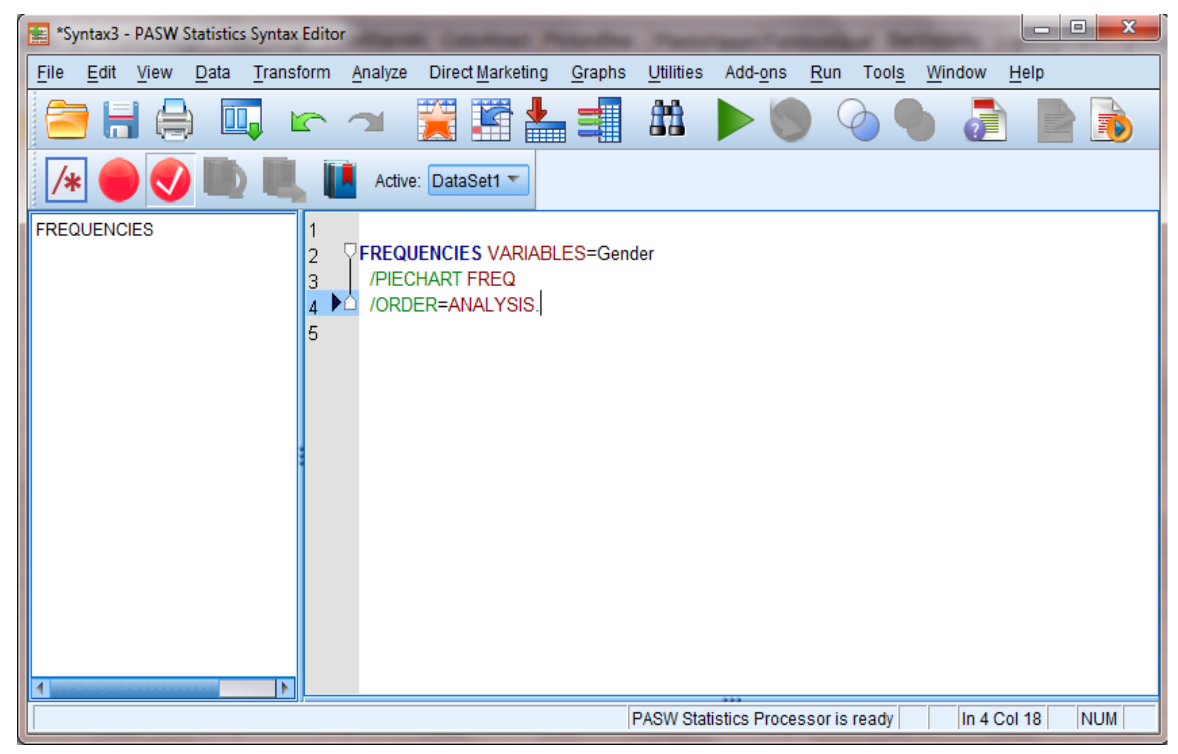
تظهر نافذة المخرجات *Output Viewer* نتائج التحليلات الإحصائية والرسومات البيانية التي يتم إجراؤها على البيانات الموجودة في نافذة محرر البيانات *Data Editor*.



الشكل 3.2. نافذة المخرجات

3-1. نافذة محرر التعليمات ***Syntax Editor***

وهي الشاشة التي يتم من خلالها كتابة التعليمات للعمليات المختلفة. ويمكن حفظ هذه التعليمات وتعديلها وتنفيذها في أي وقت. وعلى الرغم من أن معظم المهام يمكن إنجازها من خلال الاختيار والنقر بشكل مباشر إلا أن محرر التعليمات *Syntax Editor* يسمح للمستخدم بحفظ سجل للتحليلات التي قام أو يقوم بإجرائها. ويمكن تفعيل هذه النافذة من خلال النقر على زر اللصق *Paste* لدى إجراء أي تحليل أو أية عملية ضمن *SPSS*.



الشكل 4.2. نافذة محرر التعليمات

2- ملفات SPSS: *SPSS Files*

يوجد ثلاثة أنواع رئيسية من الملفات ضمن نظام *SPSS*، وهي: ملفات البيانات *Data Files* وملفات المخرجات الإحصائية *Output*وملفات التعليمات *Syntax*.

تحتوي ملفات البيانات على البيانات الخام التي يتم إدخالها من خلال نافذة محرر البيانات *Data Editor*. ويميز هذه البيانات اسمها الذي ينتهي باللاحقة *SAV*. ويتم فتح هذا النوع من الملفات من خلال شاشة عرض البيانات *Data Viewer*.

أما ملفات المخرجات الإحصائية فتحتوي على نتائج الإجراءات الإحصائية التي تظهر في شاشة المخرجات. ويتميز اسم ملف المخرجات الإحصائية بانتهائه باللاحقة *SPV*.

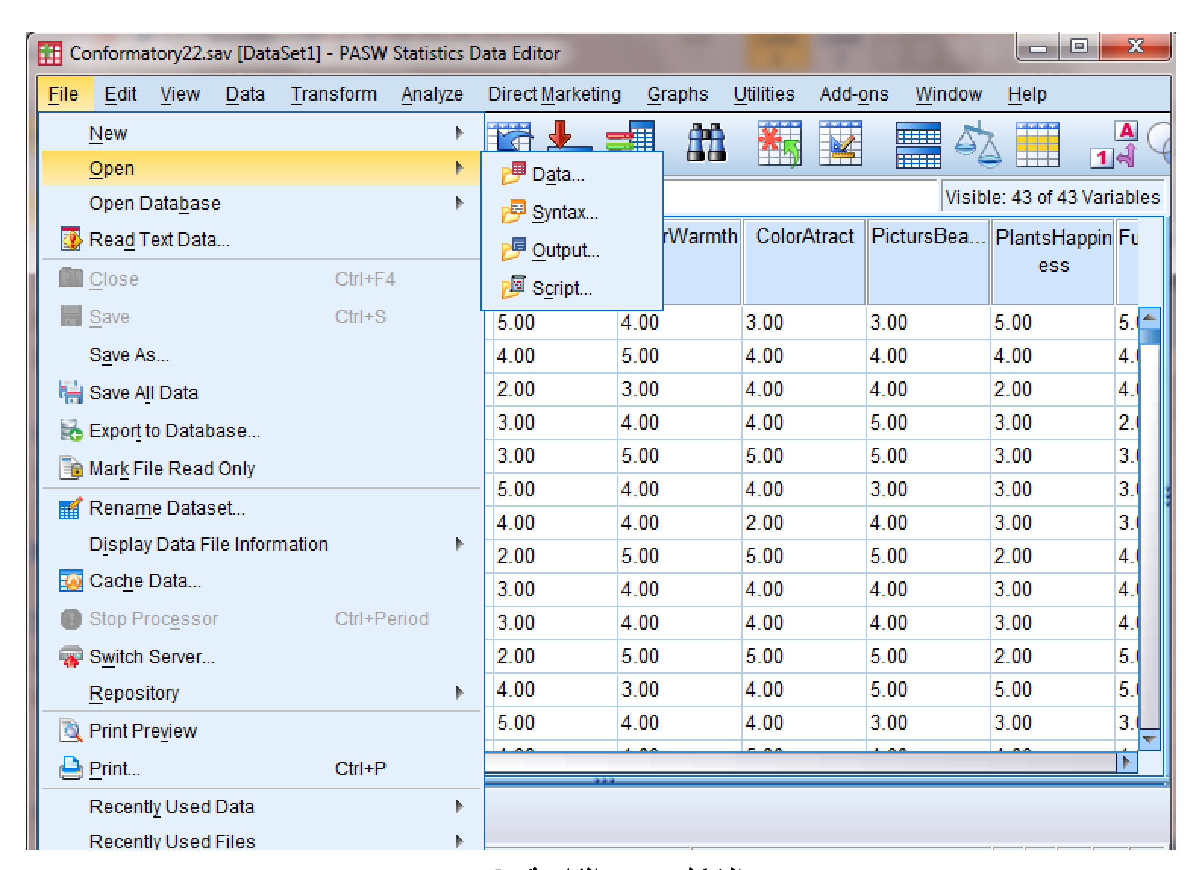
ويحتوي ملف التعليمات على التعليمات التي تم أو يتم تنفيذها كالإجراءات الإحصائية مثلاً، ويميز هذا الملف اللاحقة *SPS*.

3- قوائم SPSS: SPSS menus

يتضمن *SPSS* العديد من القوائم المتاحة للمستخدم. وتشمل هذه القوائم: الملف *"File"* و التحرير *"Edit"* والعرض *"View"* والبيانات *"Data"* والتحويل *"Transform"* والتحليل الإحصائي *"Analyze"* والأشكال *"Graphs"* والأدوات *"Utilities"* والنافذة *"Window"* والمساعدة *"Help"*. وسنورد فيمايلي شرحاً موجزاً للخيارات المتاحة ضمن كل قائمة.

3-1. قائمة الملف File Menu

يهدف استخدام هذه القائمة إلى التعامل مع الملفات من حيث: إنشاء ملفات جديدة أو فتح ملفات مخزنة أو تخزين الملفات أو طباعة الملفات، وكذلك الخروج من نظام *SPSS*. وباختصار فإن هذه القائمة تتضمن كافة الخيارات التي اعتدنا إيجادها في قائمة "الملف" في البرمجيات المختلفة.

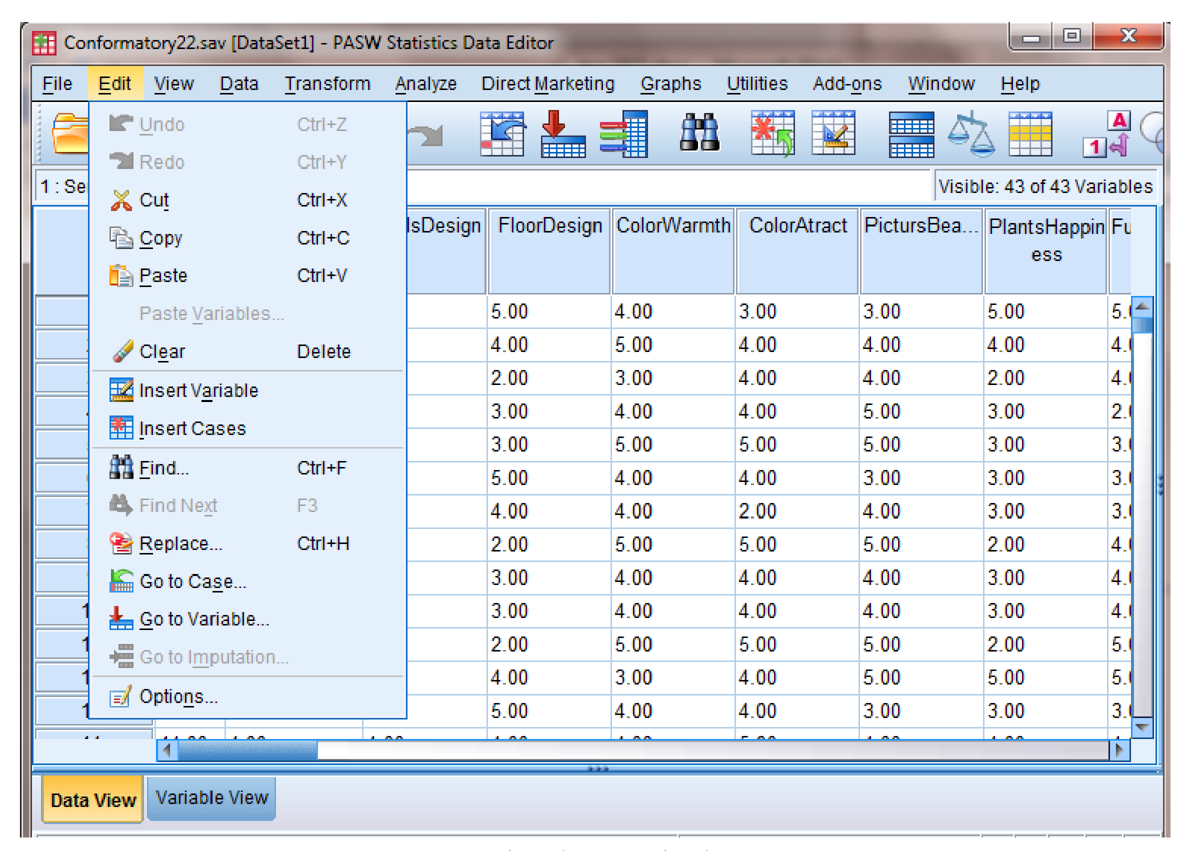


الشكل 5.2. القائمة *File*

3-2. قائمة التحرير Edit Menu

تحتوي هذه القائمة على الكثير من الأدوات المهمة مثل نسخ ونقل البيانات من مكان إلى آخر والبحث

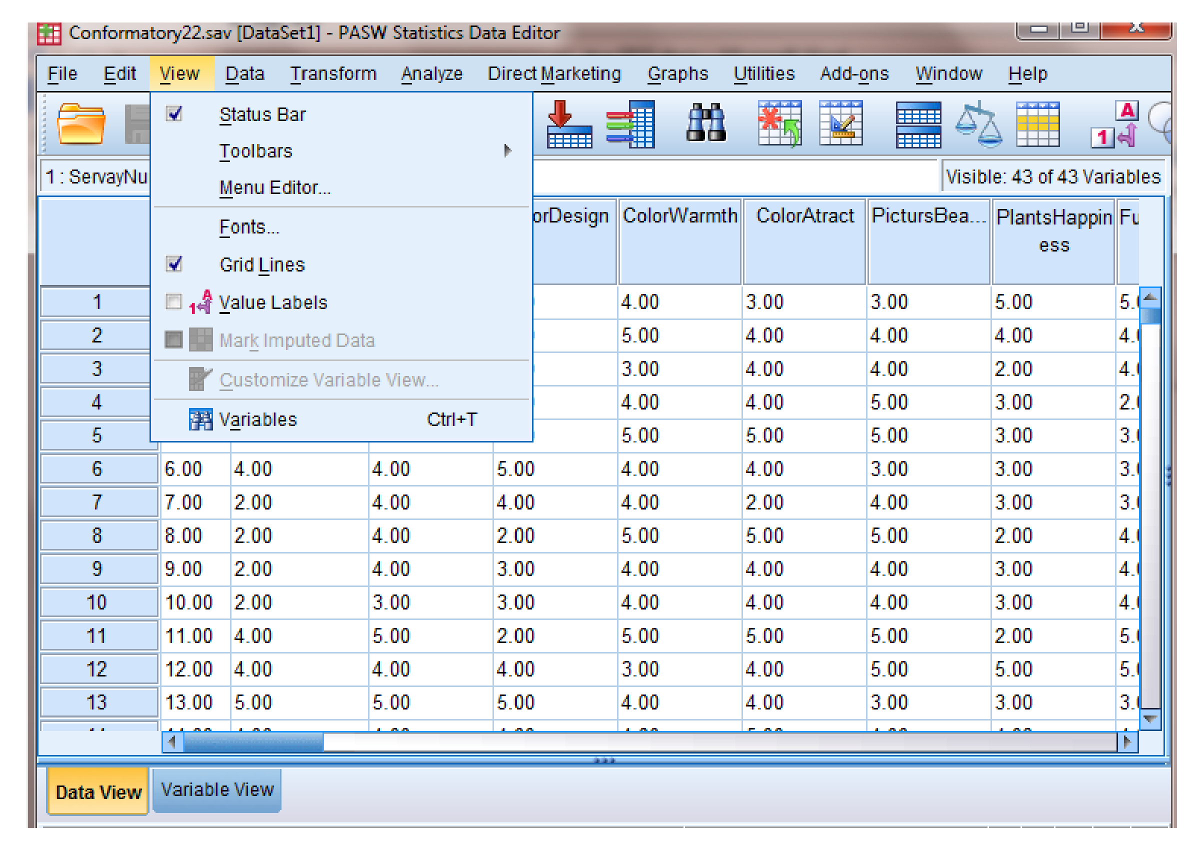
والاستبدال والعديد من الخيارات الأخرى. تتيح هذه القائمة مثلاً نسخ الأرقام أو كتل البيانات من مكان إلى آخر ضمن محرر البيانات *Data Editor*. كما يمكن استخدام الخيار *Options* لاختيار نوع الخط الذي نفضله ونمط الأرقام وغير ذلك من الخيارات المتعددة المتاحة ضمن هذه القائمة.



الشكل 6.2. القائمة Edit

3-3. قائمة العرض *View Menu*

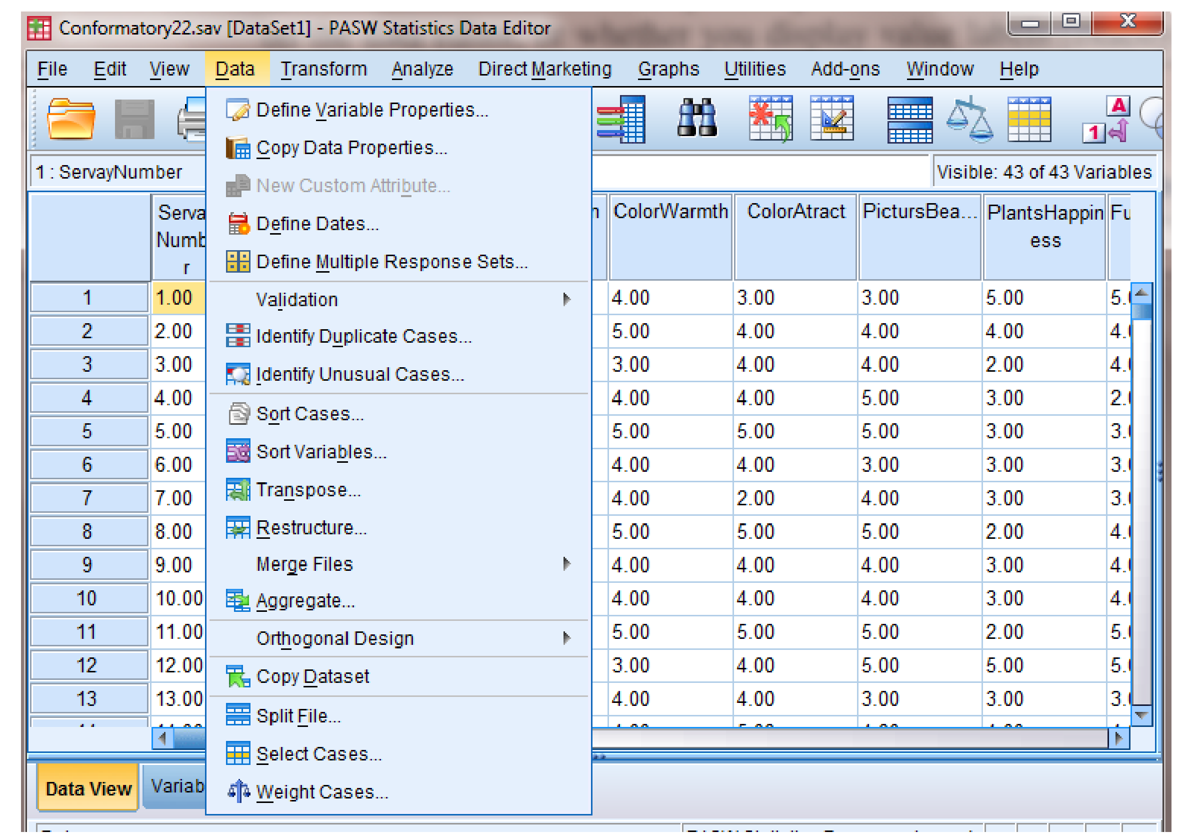
تستطيع عن طريق هذه القائمة إظهار شريط الأدوات (الأيقونات المختصرة المناسبة) *Toolbar*، وكذلك تستطيع من خلال هذه القائمة إظهار أو إخفاء خطوط الشبكة *Gridline* وتغيير نوع وحجم الخط المستخدم وإظهار أو إخفاء عناوين (دلالات) القيم *Value Label*.



الشكل 7.2. القائمة *View*

4-3. قائمة البيانات **Data Menu**

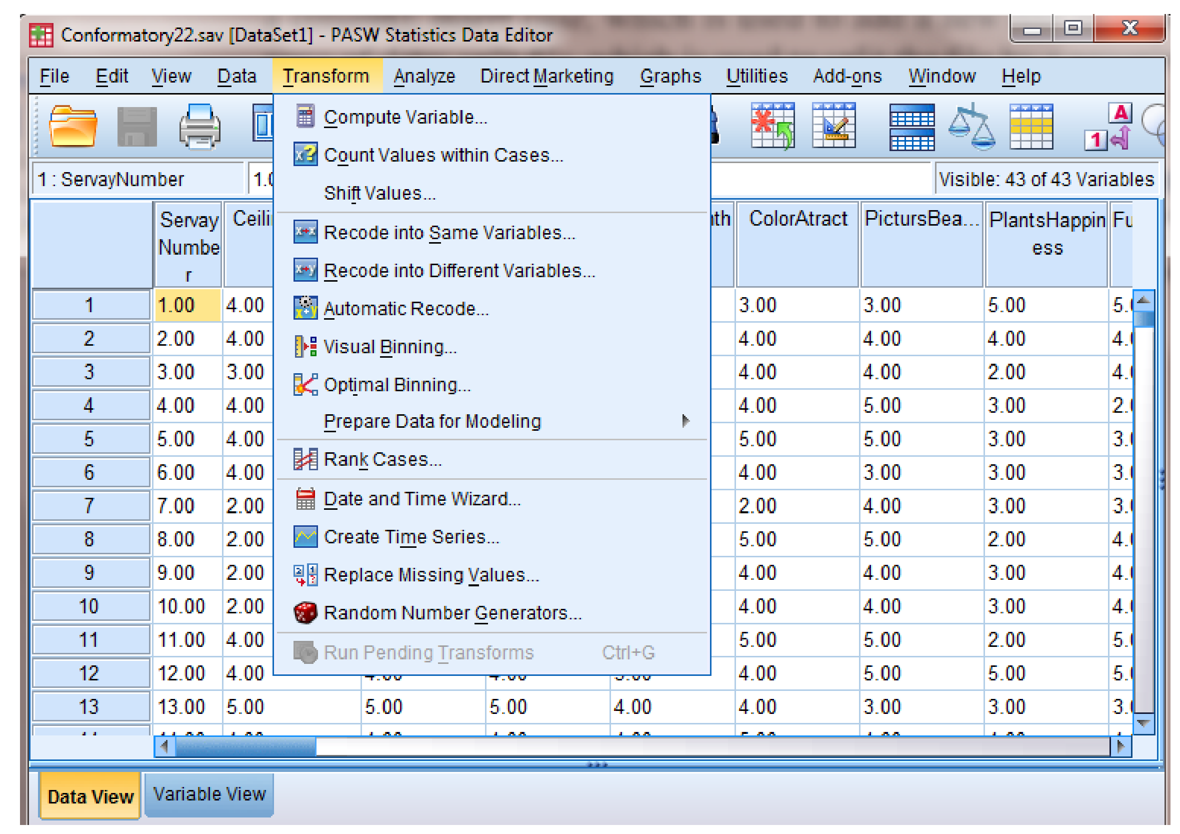
تسمح هذه القائمة بإجراء العديد من التعديلات على البيانات ضمن محرر البيانات *Data Editor*. فيمكنك مثلاً إدراج متغير جديد *Insert Variable* (تتم إضافته ضمن عمود جديد). يمكن أيضاً إدراج حالة *Insert Case* حيث تتم إضافة سطر جديد بين سطرين حاليين. كما تسمح هذه القائمة بتعريف المغيرات وتغيير أسمائها، وكذلك القيام بالعمليات المختلفة على البيانات من فرز وتحويل وتثقيل ودمج مع بيانات أخرى أو حتى تقسيم ملف البيانات...



الشكل 8.2. القائمة *Data*

3-5. قائمة التحول ***Transform Menu***

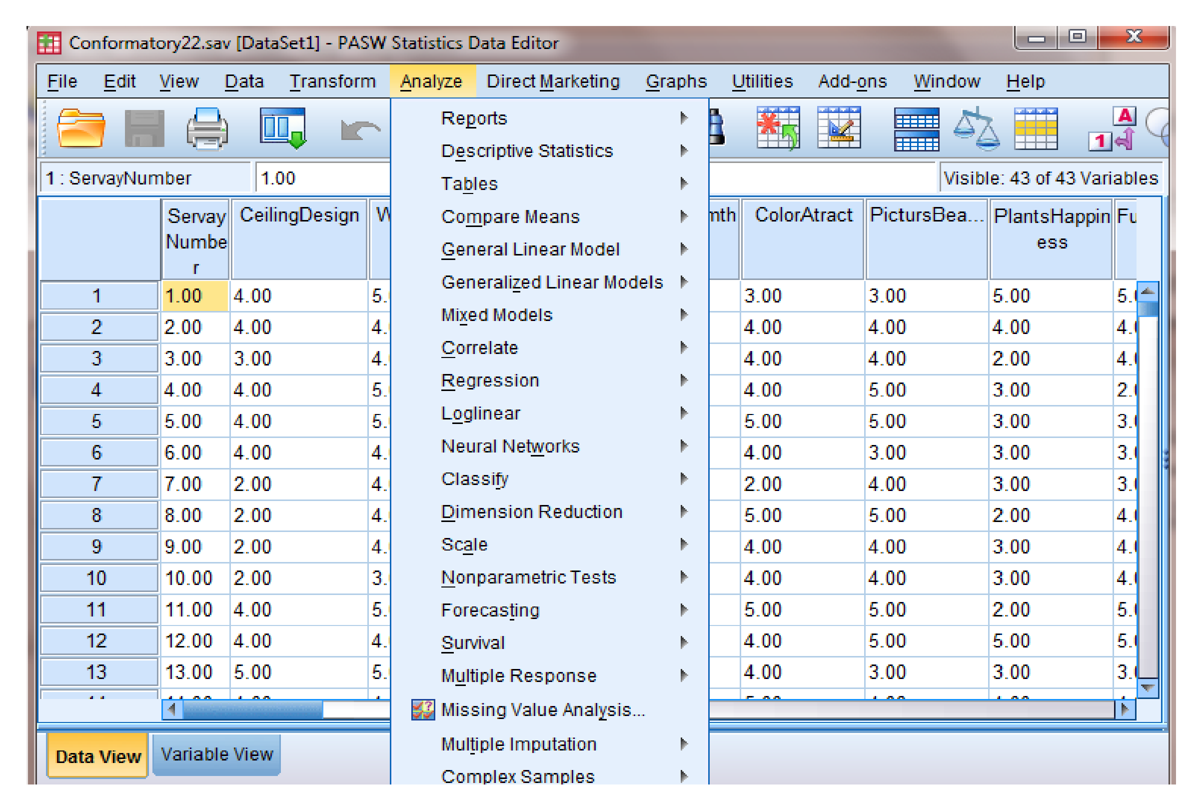
تستطيع من خلال هذه القائمة القيام بالعمليات الحسابية المختلفة وإعادة ترميز البيانات. فيمكن مثلاً من خلال هذه القائمة استخدام *recode* لتغيير قيم بعض المتغيرات. كما يفيد الخيار *compute*في تحويل البيانات أو إجراء العمليات الحسابية عليها (مثلاً: يمكن استخدام هذا الخيار لخلق متغير جديد يكون ناتجاً عن الوسط الحسابي لقيم عدة متغيرات موجودة مسبقاً).



الشكل 9.2. القائمة *Transform*

3-6. قائمة التحليل الإحصائي ***Analyze Menu***

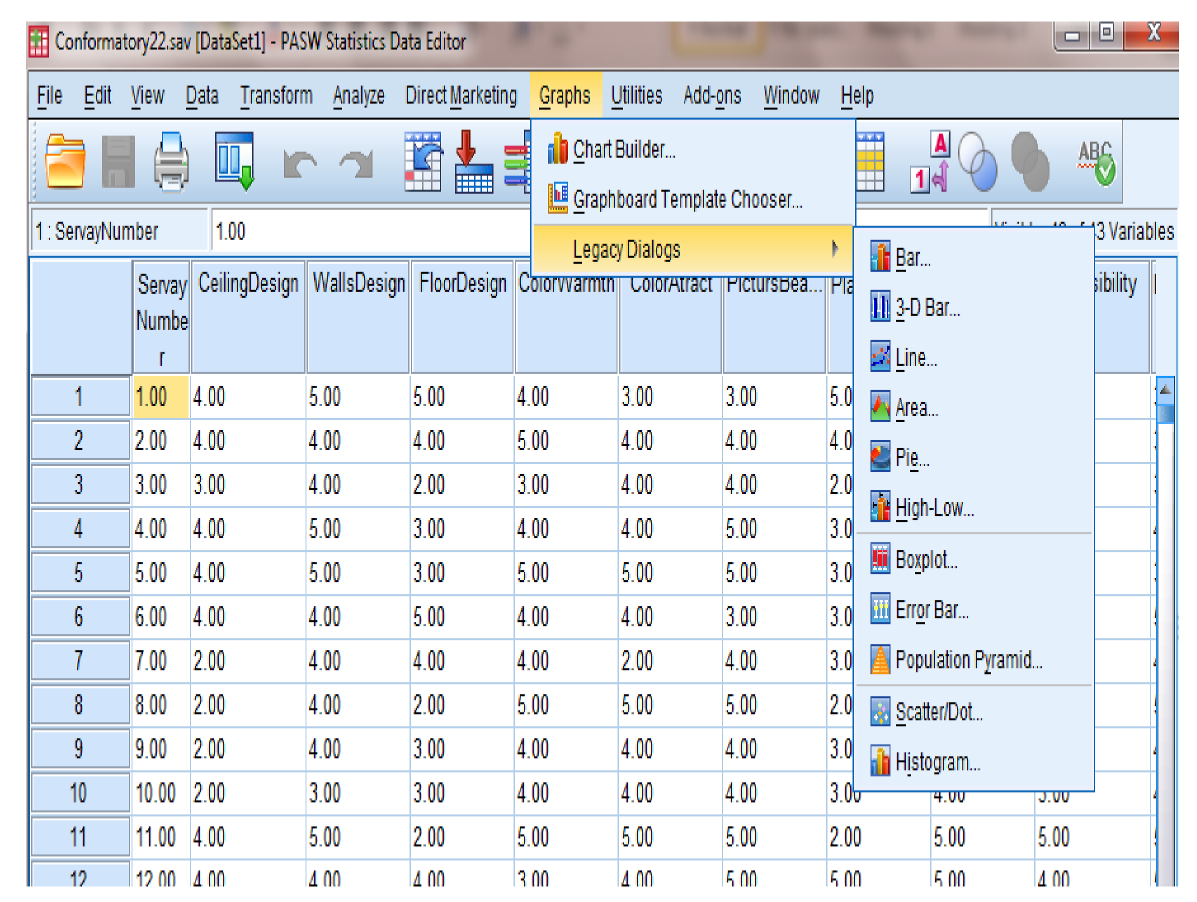
هنا تبدأ المتعة حيث تتضمن هذه القائمة التحليلات والإجراءات الإحصائية الأساسية والمتقدمة. بشكل مختصر، تتضمن هذه القائمة العديد من المؤشرات أو الإحصاءات الوصفية *Descriptive Statistics* واختبارات مقارنة المتوسطات *Compare Means* والنماذج الخطية العامة *General Linear Model* والارتباط *Correlate* والانحدار *Regression* واختصار البيانات *Data Reduction* واختبارات المقاييس *Scale* والاختبارات اللامعلمية *Nonparametric Tests* وغيرها.



الشكل 10.2. القائمة *Analyze*

3-7. قائمة الأشكال **Graphs Menu**

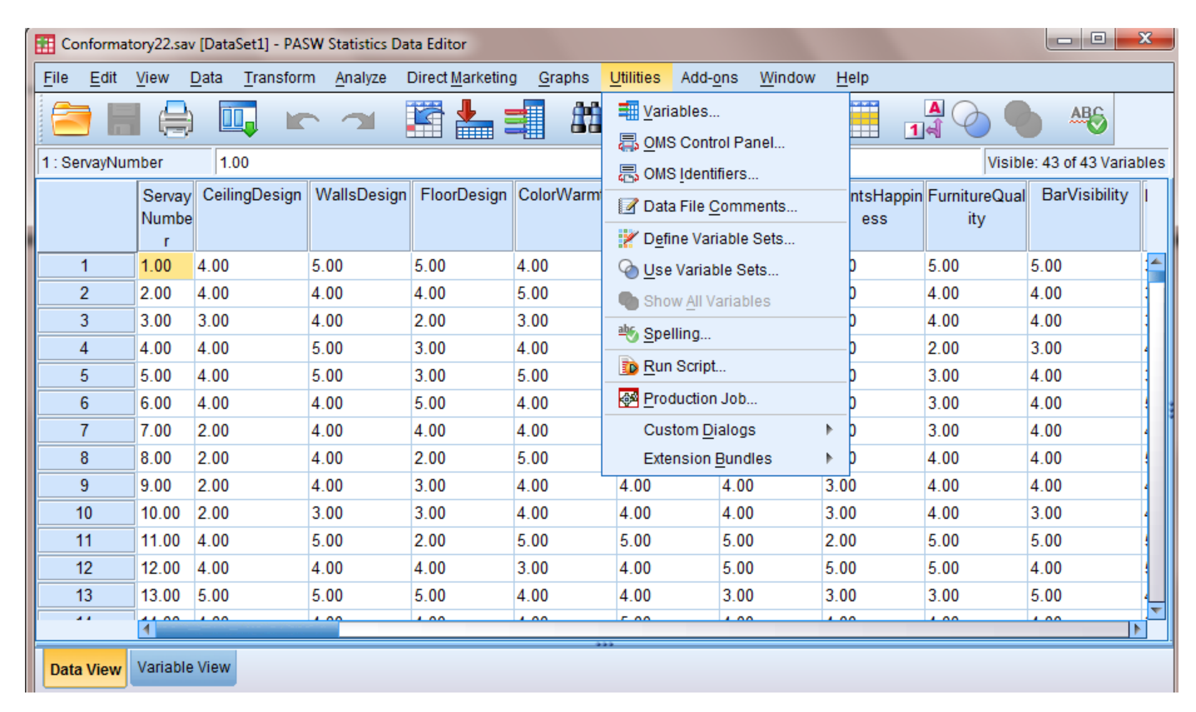
نستطيع من خلال هذه القائمة التعامل مع الرسومات البيانية وبأشكال مختلفة. وتتضمن هذه الأشكال: المدرج *histograms* والأعمدة البيانية *bar charts* وخرائط الانتشار *scatterplots* والأشكال الدائرية *pie charts* والأشكال الخطية *Line graphs* وغيرها. كما يتيح *SPSS* إمكانية تحرير الأشكال البيانية بالطريقة التي يراها المستخدم مناسبة.



الشكل 11.2. القائمة *Graphs*

3-8. قائمة الأدوات ***Utilities Menu***

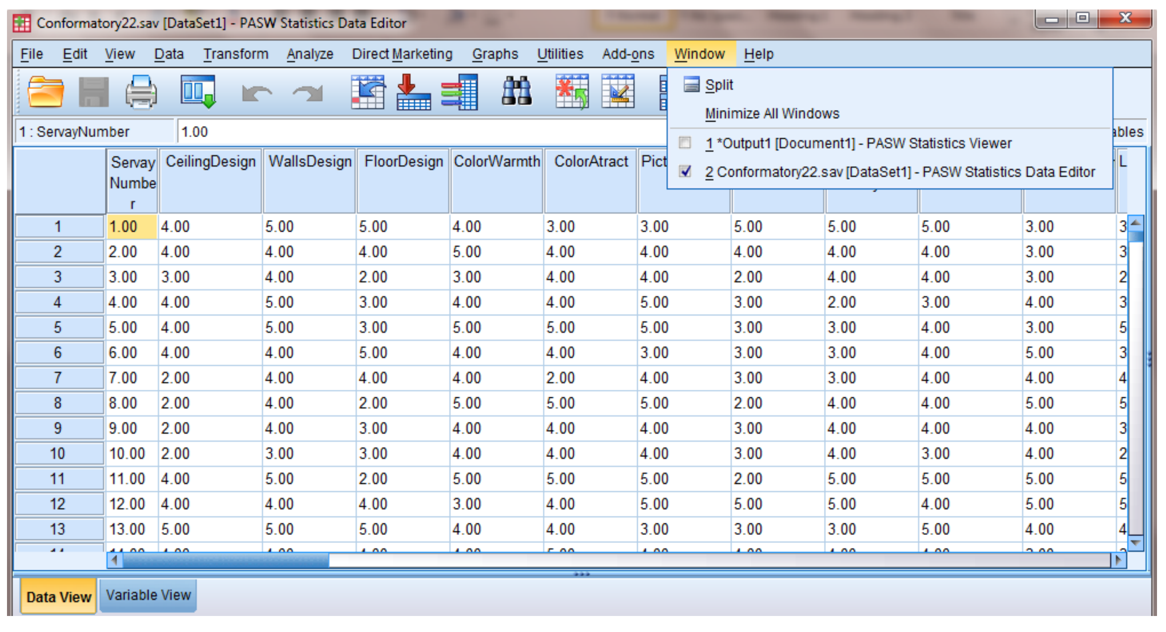
نستطيع هنا إيجاد معلومات مفصلة عن الملف المستخدم والمتغيرات التي يحويها هذا الملف، وتعريف واستخدام المجموعات *Sets* للمتغيرات المختلفة.



الشكل 12.2. القائمة *Utilities*

3-9. قائمة النافذة *Window Menu*

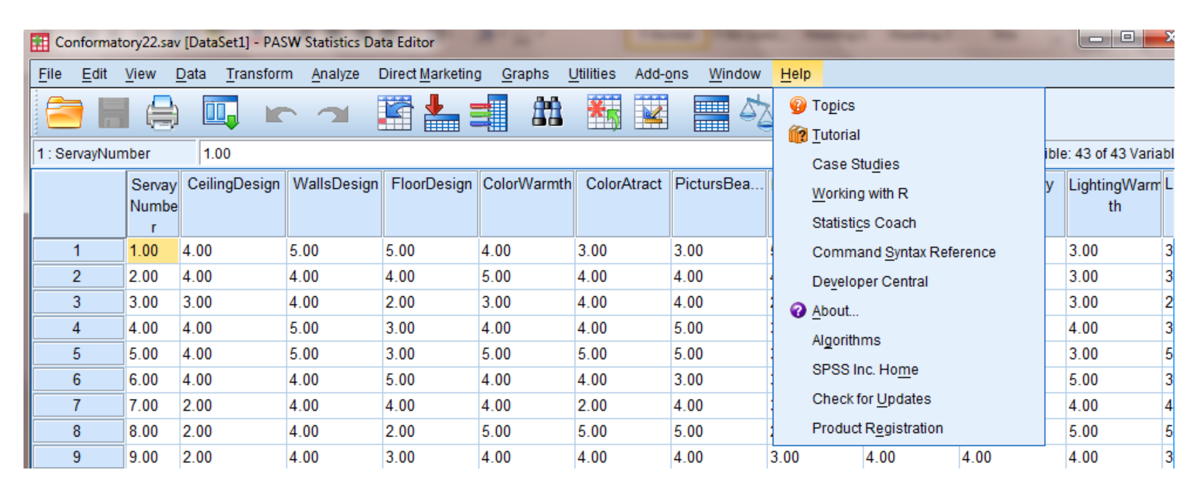
تستطيع عن طريق هذه القائمة التنقل بين النوافذ المختلفة والتحكم بحجم هذه النوافذ.



الشكل 13.2. القائمة *Window*

3-10.قائمة المساعدة Help Menu

تزودنا هذه القائمة بنظام مساعدة تفاعلي، نستطيع من خلاله الحصول على إجابات وإيضاحات للتساؤلات التي تثار عند مواجهة مشكلة ما مع نظام *SPSS*.



الشكل 14.2. القائمة *Help*

4- إدخال البيانات باستخدام SPSS: *Entering data in SPSS*

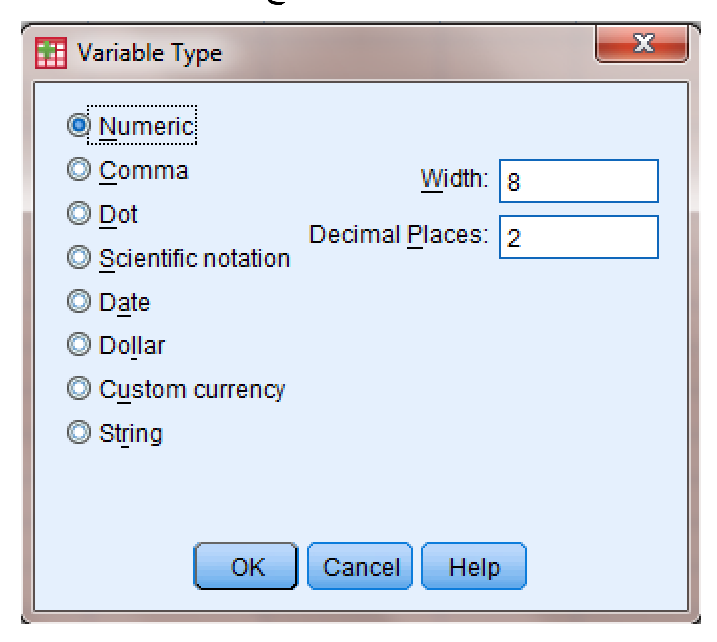
تتم عملية إدخال البيانات من خلال محرر البيانات *Data Editor*. تكون نافذة *Data View* فارغة طبعاً عند البدء بإدخال البيانات. وتتدرج عملية إدخال البيانات عادة من تعريف المتغيرات الموجودة في الاستبانة أو الدراسة وصولاً إلى إدخال إجابات أفراد العينة (عند استخدام الاستبانة) أو إدخال بيانات الدراسة. نذكّر هنا بأن كل سطر تعود بياناته لمشاهدة أو حالة معينة (شخص أو مشاهدة مثلاً) وكل عمود يخص متغيراً محدداً.

يتم تعريف المتغيرات باستخدام نافذة عرض المتغيرات *Variable View*. ويعبّر كل سطر فيها عن متغير واحد. وفيمايلي شرح مختصر للأعمدة الظاهرة في هذه النافذة:

- **اسم المتغير *Name***: وهو اسم مختصر (عادةً) يدل على المتغير. سيظهر هذا الاسم في أعلى أحد الأعمدة عند العودة إلى نافذة عرض البيانات *Data View*. ولا بد من مراعاة الأمور التالية عند كتابة اسم المتغير:

- أن يبدأ الاسم بحرف وألا ينتهي بنقطة أو "_"
- لا يتجاوز عدد الأحرف 64
- لا يجب أن يتكرر اسم المتغير
- لا يمكن استخدام الفراغ بين أحرف الاسم
- لا يمكن استخدام بعض الرموز مثل % ^ / # $ & * ؟ : ! ، " ;
- لا يمكن استخدام الأقواس
- لا يمكن استخدام كلمات مفتاحية مخصصة لأغراض معينة ضمن برنامج *SPSS*، مثل: *ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH*

- **نوع المتغير *Variable Type***: يمكن تحديد الأنواع التالية للمتغيرات:

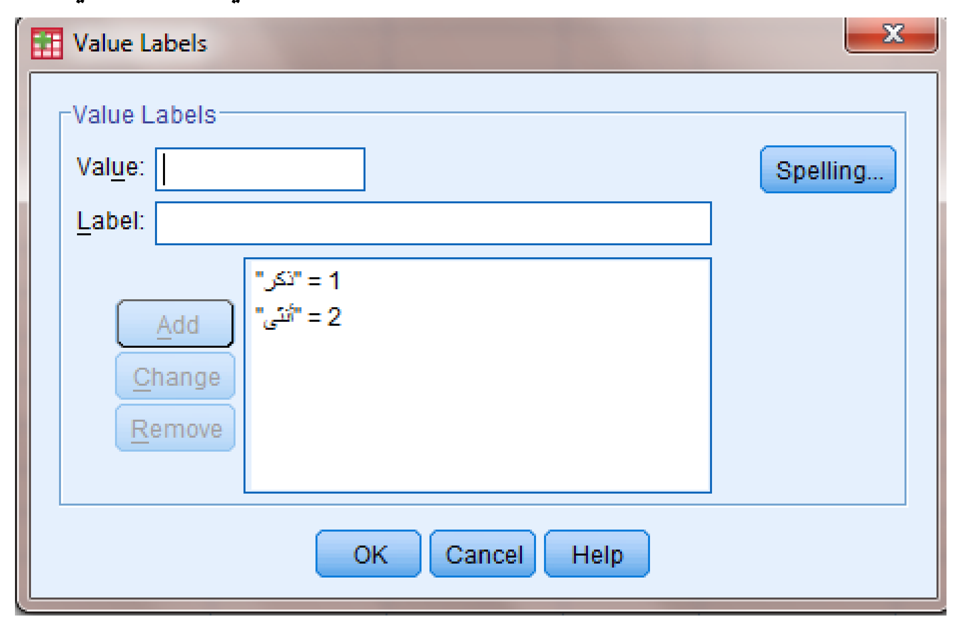


الشكل 15.2. نوع المتغير

- **العددي *Numeric***: وهو النوع الافتراضي للمتغيرات في النافذة *Variable View*.
- **الفاصلة *Comma***: وهو متغير عددي مع إضافة فاصلة (,) للفصل بين كل ثلاث مراتب صحيحة. مثلاً العدد 881556223 يكتب *881,556,223* بموجب هذا النوع. وتستعمل النقطة للأرقام العشرية.

- **النقطة *Dot***: وهو متغير عددي مع استخدام (.) لفصل كل ثلاث مراتب صحيحة، فالعدد أعلاه يكتب *881.556.223* بموجب هذا النوع. وتستعمل الفاصلة للفصل بين الجزء الصحيح والجزء العشري للأرقام العشرية.
- **التاريخ *Date***: متغير يمثل التاريخ أو الوقت.
- **الدولار *Dollar***: يستعمل كرمز للدولار.
- **العملة *custom Currency***: متغير يحدد من قبل المستفيد للدلالة على العملة المطلوبة.
- **النص *string***: وهو متغير تكون بياناته على شكل أحرف أو كلمات أو أرقام بلا دلالة كمية.

- **العرض *Width***: يبين عدد مراتب المتغير كعدد حروف النص.
- **المنازل العشرية *Decimals***: عدد الخانات العشرية (بعد الفاصلة).
- **التوصيف *Label***: يستخدم لتوصيف المتغير. لاحظ أن وضع المؤشر فوق اسم المتغير في النافذة *Data View* سيؤدي إلى ظهور التوصيف.
- **القيم *Values***: يستخدم لتحديد معنى الأرقام أو الرموز المستخدمة للبيانات. فمثلاً عند إدخال متغير الجنس يمكن إعطاء الرمز "1" للذكور و "2" للإناث كما في الشكل التالي.

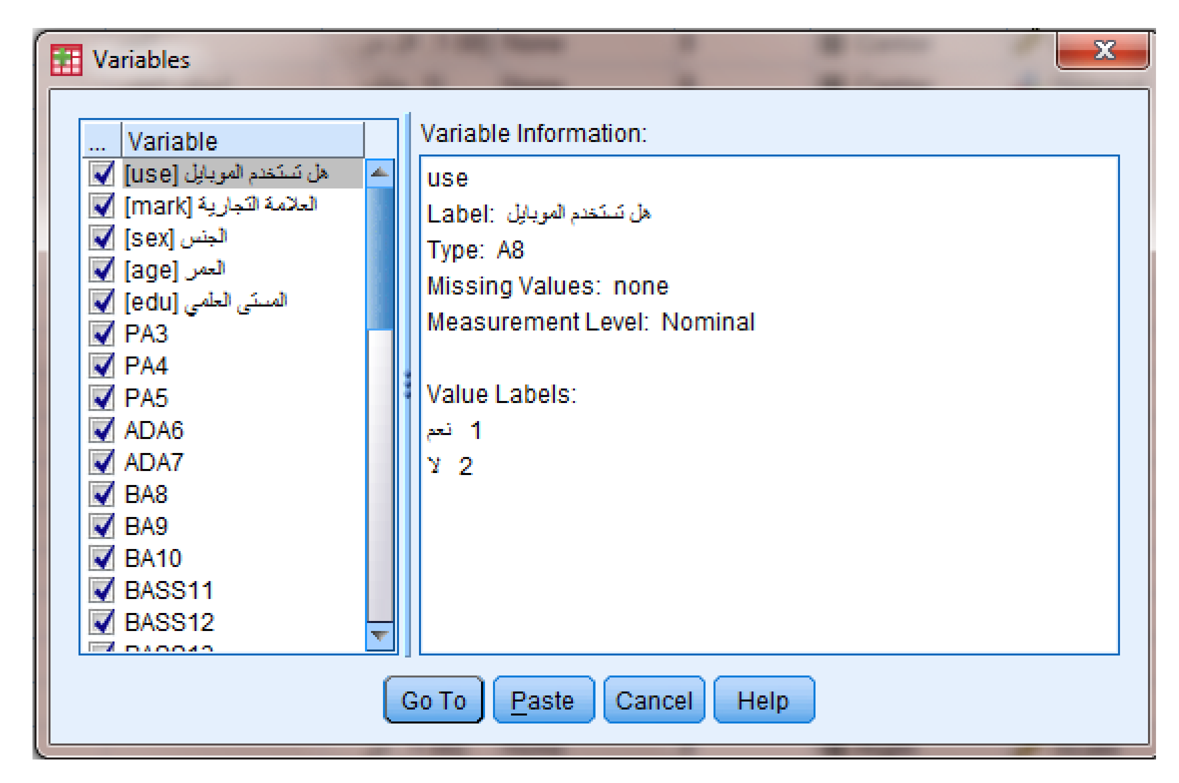


الشكل 16.2. نافذة قيم ترميز المتغير

- **القيم المفقودة *Missing***: يمكن هنا تحديد رمز أو رقم للدلالة على القيم المفقودة كالرقم 9 مثلاً. ويوجد نوعان من القيم المفقودة في *SPSS*.
  - النوع الأول: هي القيم المفقودة التي تحدد من قبل المستخدم. ويتم تعريفها بواسطة شاشة الحوار *Missing Values*
  - النوع الثاني: هي قيم المتغير المفقودة أصلاً، أي أنها خلايا فارغة، نتيجة عدم الاستجابة من قبل بعض الأشخاص لسؤال معين في استبيان ما. وفي هذه الحالة فإن الخلايا الفارغة تحول تلقائياً إلى قيم مفقودة للنظام *System missing values* وهذا ينطبق على المتغيرات العددية، أما بالنسبة للمتغيرات النصية *String Variables* فإن الخلايا الفارغة تعتبر صحيحة أي أنها لا تعتبر قيماً مفقودة.

- **الأعمدة *Columns***: يحدد عرض العمود الذي يوجد فيه المتغير في نافذة عرض البيانات *Data View*.
- **المحاذاة *Align***: تحدد طريقة محاذاة البيانات (يمين، يسار، وسط) في العمود الذي تتواجد فيه بيانات المتغير في النافذة *Variable View*.
- **المقياس *Measure***: لتحديد نوع البيانات (قياسي أو مستمر*Scale*، ترتيبي *Ordinal*، اسمي *Nominal*). ويعتمد الخيار الافتراضي للمقياس ضمن *SPSS* على نوع البيانات. فمثلاً لبيانات من النوع العددي *numeric* يكون المقياس الافتراضي هو القياسي *scale* وهو يعود عادة للمقاييس المدرجة أو الفئوية ومقاييس النسب. أما لمتغيرات من نوع نصي *string* فإن الخيار الافتراضي للمقياس ضمن *SPSS* يكون المقياس الاسمي *nominal*. أما النوع الثالث أي الترتيبي *ordinal* فلا يظهر كخيار افتراضي. ويمكن طبعاً تغيير الخيار بسهولة من قبل المستخدم.

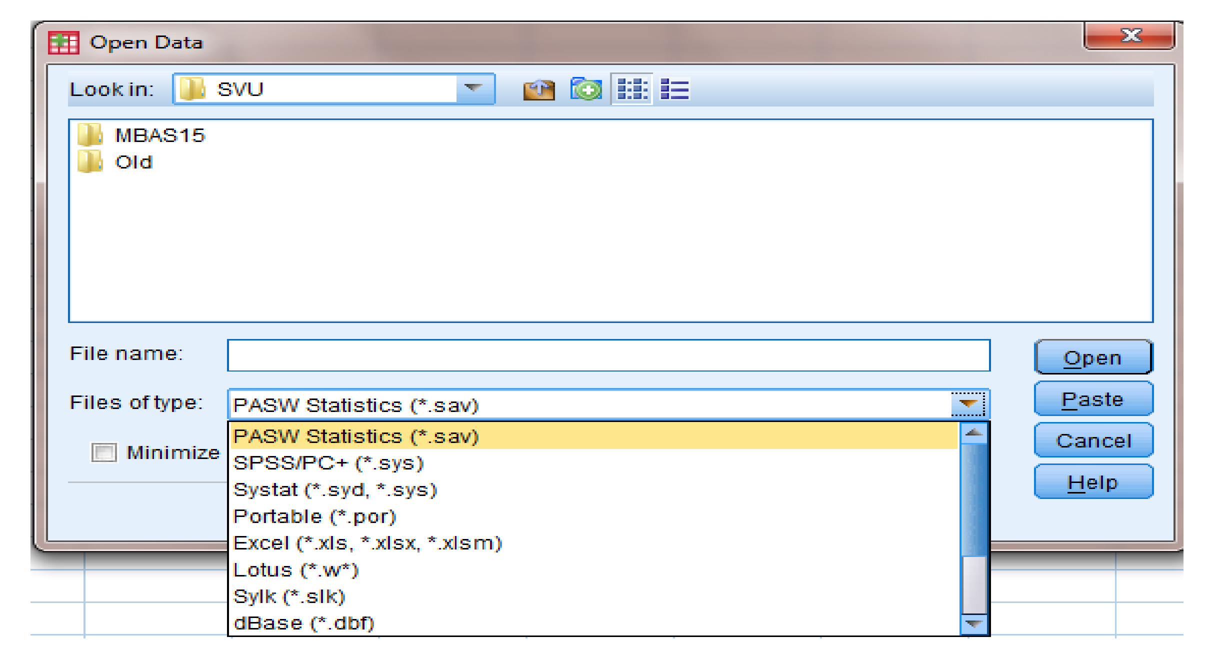
بعد الانتهاء من تعريف المتغيرات يمكن الحصول على ملخص لخصائص كل متغير من خلال الخيار *Variables*ضمن قائمة الأدوات *Utilities*.



الشكل 17.2. نافذة خصائص المتغيرات

يتم إدخال البيانات الخاصة بالدراسة ضمن النافذة *Data View*. وتشبه عملية إدخال البيانات ضمن الجدول في *Data View* عملية إدخال البيانات ضمن أي من البرامج التي تعنى بجداول البيانات مثل *Excel*.

وفي النهاية لا بد من الإشارة إلى أنه بإمكان المستخدم استيراد بيانات مخزنة ضمن ملفات تم إنشاؤها باستخدام برامج أخرى غير *SPSS*. ويتم ذلك باتباع المراحل التالية *Data*←*Open*←*File* واختيار نوع الملف المراد استيراده أو فتحه ضمن القائمة المنسدلة *Files of type* في النافذة *Open Data*.



الشكل 18.2. فتح ملفات البيانات ضمن *SPSS*

قائمة المراجع

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام *SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Coakes S.J. (2005), **SPSS for Windows: Analysis without Anguish**, John Wiley, Australia.*
- *Field A. (2006), **Discovering Statistics Using SPSS**, 2nd Edition, SAGE, England.*
- *Landau S. &Everitt B.S. (2004), **A Handbook of Statistical Analysis Using SPSS**, Chapman & Hall/CRC Press, USA.*
- *Pallant J. (2007), **SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows**, 3dEdition, McGraw Hill, USA.*
- ***SPSS Statistics Base 17.0 User's Guide (2007**), SPSS Inc., USA.*

مقترحات وتمارين للفصل الثاني

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1- عدد أنواع النوافذ المتاحة في *SPSS*؟

(الحل في الفقرة: 1.)

2- ما هي أنواع الملفات المتاحة في *SPSS*؟

(الحل في الفقرة: 2.)

3- ما هي النافذة في *SPSS* التي يتم من خلالها عرض وتعريف المتغيرات؟

(الحل في الفقرة: 4.)

4- قم بإنشاء ملف *SPSS* وتعريف المتغيرات التالية:

الجنس: 1. ذكر 2. أنثى

العمر: عدد السنوات

الحالة الاجتماعية: 1. عازب 2. متزوج 3. أرمل 4. مطلق

العمل: 1. عاطل عن العمل 2. موظف قطاع عام 3. موظف قطاع خاص 4. أعمال حرة

الدخل الذي احصل عليه يغطي احتياجاتي الأساسية

غير موافق بشدة 2. غير موافق 3. حيادي 4. موافق 5. موافق بشدة

يتوجب على الفرد القيام بأكثر من عمل لتلبية احتياجاته الاساسية

غير موافق بشدة 2. غير موافق 3. حيادي 4. موافق 5. موافق بشدة

(الحل في الفقرة: 4.)

الفصل الثالث التحليل الوصفي

Chapter 3: Descriptive analysis

الكلمات المفتاحية:

التكرارات، مقاييس النزعة المركزية، مقاييس التشتت، التوزيع الطبيعي.

ملخص:

يشرح الفصل كيفية تطبيق التحليل الوصفي في *SPSS*. ويبين كيفية الحصول على التكرارت والوسط الحسابي والوسيط والمنوال والانحراف المعياري والتباين. كما يتطرق الفصل إلى كيفية التعرف على شكل التوزيع الطبيعي من خلال المدرج التكراري ومقاييس الالتواء والتفلطح.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- فهم كيفية وصف العينة باستخدام التحليل الوصفي
- التعرف على مقاييس النزعة المركزية
- الإلمام بمقاييس التشتت
- استيعاب كيفية التعرف على شكل التوزيع الطبيعي

الفصل الثالث: التحليل الوصفي

يستخدم الإحصاء الوصفي لسبر البيانات التي تم جمعها ولتلخيص وتوصيف هذه البيانات. ولا تحتاج عمليات عرض ووصف البيانات لمجهود كبير في *SPSS* حيث يمكن لأي مستخدم مهما كانت خلفيته تعلم الأوامر الضروية وتنفيذها. ونظراً لبساطتها وسهولة فهمها تستخدم الإحصاءات الوصفية على نطاق واسع من قبل الباحثين.

إذاً، بعد التأكد من صحة البيانات تبدأ عملية تحليل البيانات بالتحليل الوصفي عادةً. وتستعمل الإحصاءات الوصفية لاغراض متعددة، نذكر منها:

- وصف خصائص العينة
- والتأكد من احترام المتغيرات لشروط طرق التحليل الإحصائي التي ستستخدم لاحقاً في التحقق من تساؤلات وفرضيات البحث
- وللإجابة على بعض التساؤلات المحددة.

وكما ذكرنا في فصل سابق فإن نوع المقياس (مستوى القياس) يمكن أن يساعد الباحث في تحديد الأداة الإحصائية الأنسب. نذكر أيضاً بأن الإحصاءات التي يمكن تطبيقها في حالة المقاييس الأقل غنى بالمعلومة (كالمقاييس الاسمية) يمكن تطبيقها أيضاً بشكل صحيح مع المقاييس الأغنى بالمعلومة (مقاييس النسب). وتعتبر الأشكال البيانية وسيلة مفيدة وضرورية أحياناً لوصف البيانات.

1- التكرارات *Frequencies*

يستخدم التكرار *Frequency* عادةً للحصول على عدد الإجابات لكل حالة من حالات المتغير أو لوصف توزيع مفردات العينة حسب نوع المتغير وحالاته كما يستخدم للحصول على النسب المئوية لهذه الأعداد. يبين التكرار إذاً عدد مرات حصول أو ظهور كل قيمة في ملف البيانات. ويمكن إظهار التكرارات بشكل جدولي أو بشكل بياني/غرافيكي.

سنستخدم في هذا الفصل بيانات تم جمعها من خلال استبانة تم تصميمها من قبل أحد الباحثين لشرح استخدامات الإنترنت للأغراض الشخصية أو غير المهنية. يظهر الجدول التالي البيانات التي تم جمعها من 30 مجيباً. تضمنت الدراسة المتغيرات التالية:

- الجنس: (ذكر، أنثى)
- الألفة بالانترنت: مقياس من 7 درجات (1= غير مألوفة أبداً / 7= مألوفة جداً)
- ساعات استخدام الانترنت أسبوعياً
- الاتجاه نحو الانترنت: مقياس من 7 درجات (1= سلبي جداً / 7= إيجابي جداً)
- والاتجاه نحو التكنولوجيا: مقياس من 7 درجات (1= سلبي جداً / 7= إيجابي جداً)
- التسوق عبر الانترنت: (نعم/لا)
- الصيرفة عبر الانترنت: (نعم/لا)

الجدول 1.3. بيانات دراسة استخدام الانترنت للأغراض الشخصية

Respondent Number	Sex	Familiarity	Internet Usage	Attitude Toward Internet	Attitude Toward Technology	Usage of Internet Shopping	Usage of Internet Banking
1	1.00	7.00	14.00	7.00	6.00	1.00	1.00
2	2.00	2.00	2.00	3.00	3.00	2.00	2.00
3	2.00	3.00	3.00	4.00	3.00	1.00	2.00
4	2.00	3.00	3.00	7.00	5.00	1.00	2.00
5	1.00	7.00	13.00	7.00	7.00	1.00	1.00
6	2.00	4.00	6.00	5.00	4.00	1.00	2.00
7	2.00	2.00	2.00	4.00	5.00	2.00	2.00
8	2.00	3.00	6.00	5.00	4.00	2.00	2.00
9	2.00	3.00	6.00	6.00	4.00	1.00	2.00
10	1.00	9.00	15.00	7.00	6.00	1.00	2.00
11	2.00	4.00	3.00	4.00	3.00	2.00	2.00
12	2.00	5.00	4.00	6.00	4.00	2.00	2.00
13	1.00	6.00	9.00	6.00	5.00	2.00	1.00
14	1.00	6.00	8.00	3.00	2.00	2.00	2.00
15	1.00	6.00	5.00	5.00	4.00	1.00	2.00
16	2.00	4.00	3.00	4.00	3.00	2.00	2.00
17	1.00	6.00	9.00	5.00	3.00	1.00	1.00
18	1.00	4.00	4.00	5.00	4.00	1.00	2.00
19	1.00	7.00	14.00	6.00	6.00	1.00	1.00
20	2.00	6.00	6.00	6.00	4.00	2.00	2.00
21	1.00	6.00	9.00	4.00	2.00	2.00	2.00
22	1.00	5.00	5.00	5.00	4.00	2.00	1.00
23	2.00	3.00	2.00	4.00	2.00	2.00	2.00
24	1.00	7.00	15.00	6.00	6.00	1.00	1.00
25	2.00	6.00	6.00	5.00	3.00	1.00	2.00
26	1.00	6.00	13.00	6.00	6.00	1.00	1.00
27	2.00	5.00	4.00	5.00	5.00	1.00	1.00
28	2.00	4.00	2.00	3.00	2.00	2.00	2.00
29	1.00	4.00	4.00	5.00	3.00	1.00	2.00
30	1.00	3.00	3.00	7.00	5.00	1.00	2.00

تظهر تكرارات قيم المتغير ضمن *SPSS* ببساطة كبيرة ضمن ما يسمى بالجداول التكرارية والتي يمكن الحصول عليها باتباع الخطوات التالية:

Analyze←Descriptive Statistics←Frequencies

لتظهر النافذة التالية والتي يتم من خلالها اختيار المتغير أو المتغيرات المراد الحصول على جداولها التكرارية.



الشكل 1.3. نافذة *Frequencies* ضمن *SPSS*

فإذا ما أردنا توصيف متغير "الجنس" أي معرفة عدد الذكور والإناث ضمن العينة فما علينا سوى سحب المتغير "الجنس" من الإطار الموجود في جهة اليسار إلى الإطار الموجود في جهة اليمين تحت عنوان *Variable(s)*. وعند النقر على *OK* سيظهر الجدول التكراري التالي.

الجدول 2.3. نتيجة التكرارات لمتغير "الجنس"

Statistics

الجنس

N	*Valid*	*30*
	Missing	*0*

الجنس

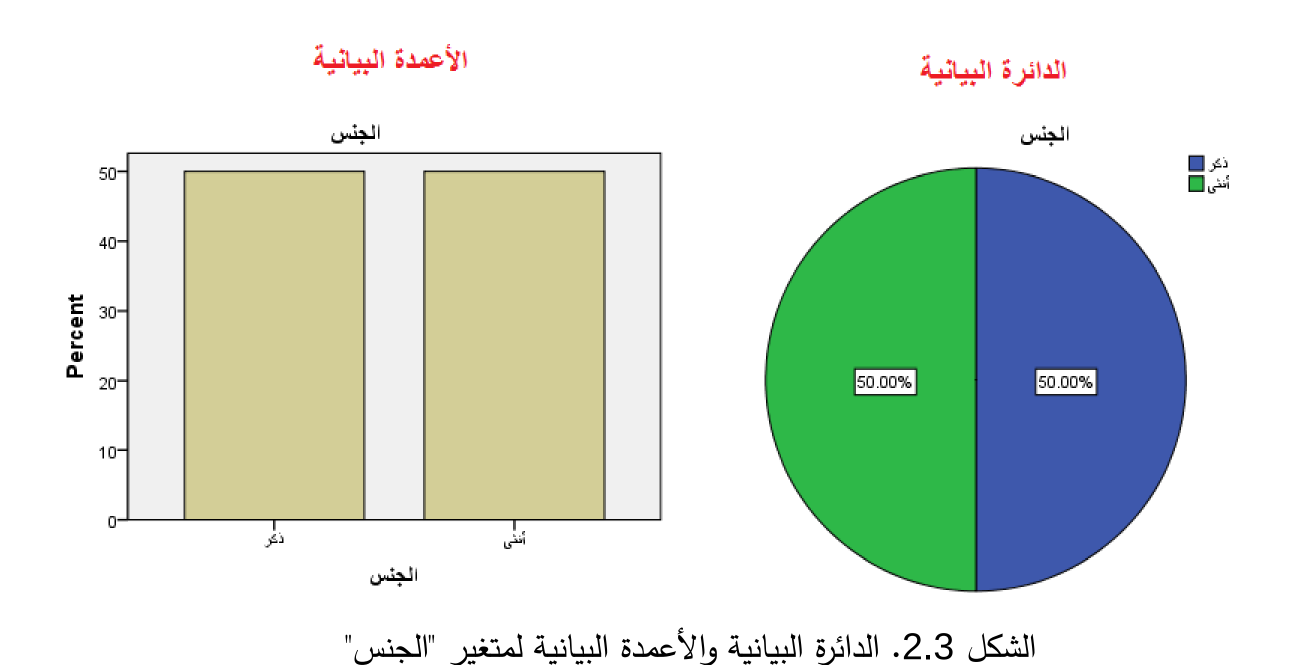
		Frequency	*Percent*	*Valid Percent*	*Cumulative Percent*
Valid	ذكر	*15*	*50.0*	*50.0*	*50.0*
	أنثى	*15*	*50.0*	*50.0*	*100.0*
	Total	*30*	*100.0*	*100.0*	

يتكون كل جدول تكراري من أربعة أعمدة:

- العمود الأول *Frequency*: يتضمن إحصاء أفراد العينة في كل فئة من فئات المتغير
- العمود الثاني *Percent*: يتضمن النسب المئوية لكل فئة.
- العمود الثالث *Valid Percent*: يتضمن النسب المئوية بعد استبعاد البيانات المفقودة.
- العمود الرابع *Cumulative Percent*: يمثل النسب التراكمية لفئات المتغير.

نلاحظ أيضاً أن النافذة *Frequencies* تتيح لنا تمثيل أو عرض التكرارات باستخدام الأشكال أو الخرائط البيانية *Charts*. ومن أهم الأشكال البيانية المتاحة نجد:

- **الدائرة البيانية *Pie Chart***: تستخدم لتمثيل التكرارات أو النسب المئوية في حالة المتغيرات النوعية أو الترتيبية.
- **الأعمدة البيانية *Bar Chart***: تستخدم لتمثيل التكرارات أو النسب المئوية في حالة المتغيرات النوعية أو الترتيبية. يشير ارتفاع كل عمود إلى تكرار القيمة.



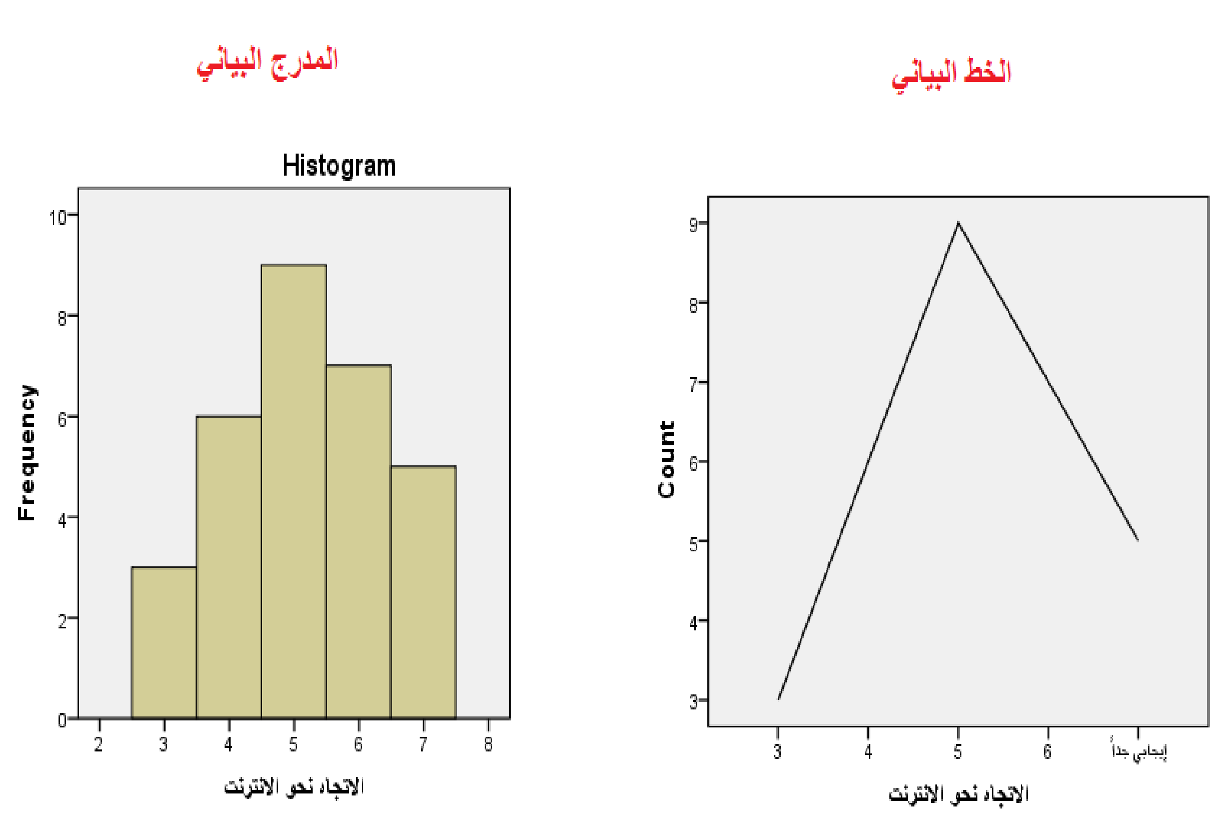
الشكل 2.3. الدائرة البيانية والأعمدة البيانية لمتغير "الجنس"

- **المدرج التكراري *Histograms***: يستخدم فقط في حالة المتغيرات القياسية أو المستمرة (المدرجة والنسب). ويفيد المدرج التكراري بشكل خاص في إظهار توزيع المتغير.

ويمكن أن يفيد **الخط البياني** ***Line Chart*** في إيضاح القيم المختلفة للمتغير حيث يصل بين هذه القيم مما يعطي انطباعاً بأن المتغير مستمر. يعمل الخط البياني بشكل جيد في حال أردنا إظهار تغيرات قيم المتغير عبر الزمن. وللحصول على الخط البياني نتبع المسار التالي:

Graphs←Legacy Dialogs←Line←Simple

نلاحظ هنا إمكانية اتباع مسار مشابه للحصول على الأشكال والخرائط البيانية السابقة (الدائرة والأعمدة البيانيةوالمدرج التكراري) وغيرها.



الشكل 3.3. المدرج البياني والخط البياني لمتغير "الاتجاه نحو الانترنت"

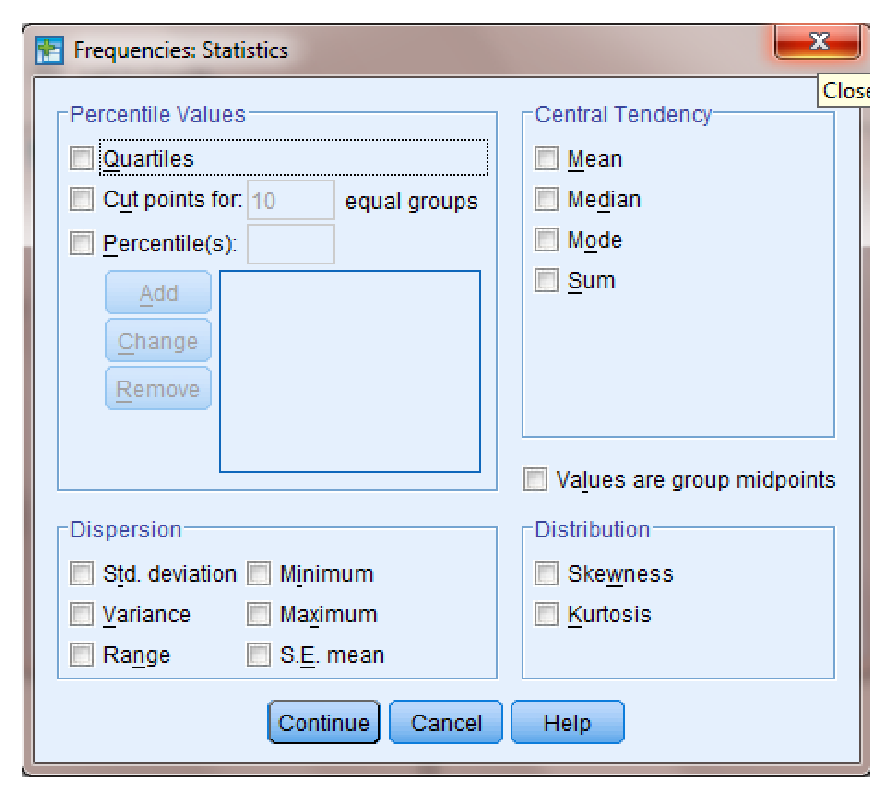
كما تتيح لنا النافذة *Frequencies* أن نختار الزر *Statistics* مما يمكننا من الحصول على مقاييس النزعة المركزية والتشتت والتوزيع والربيعيات.

2- مقاييس النزعة المركزية *Measures of central tendency*

يتضمن الإحصاء الوصفي المرتبط بمتغير ما نوعين أساسيين من المقاييس: يتعلق النوع الأول بتوصيف النزعة المركزية فيما يصف النوع الثاني التشتت.

تتضمن مقاييس النزعة المركزية *Centrality-based measures* بشكل رئيسي كلاً من الوسط الحسابي والوسيط والمنوال. وكما أشرنا سابقاً يمكن الحصول على هذه المقاييس في *SPSS* باتباع المسار التالي:

Analyze←Descriptive Statistics←Frequencies←Statistics



الشكل 4.3. نافذة *Statistics* ضمن النافذة *Frequencies*

2-1. الوسط الحسابي *Mean*

يعتبر الوسط الحسابي *Mean* أحد مقاييس الوصف الإحصائي التي يلجأ الباحث إلى استخدامها عندما يكون راغباً في إيجاد قيم مركزية معينة تستقطب حولها التوزيعات التكرارية لمشاهدات العينة.

ويحسب الوسط الحسابي عن طريق جمع كل قيم المشاهدات الخاصة بالمتغير في إطار العينة، ثم قسمة هذا المجموع على عدد المشاهدات.

$$/ n\bar{X} = \sum x$$

يعتبر الوسط الحسابي أكثر مقاييس النزعة المركزية استخداماً. وهو يستخدم مع المقاييس المستمرة أو القياسية (مدرجة ونسب). ومن أهم مزاياه سهولة الحساب والفهم بالإضافة إلى كونه يأخذ في الاعتبار جميع القيم. أما عيوبه فتتلخص في كونه يتأثر بالقيم الشاذة.

2-2. الوسيط *Median*

يعرف الوسيط ***Median*** بأنه القيمة التي تتوسط جميع القيم المعطاة عن المتغير، بحيث يقع نصف عدد تلك القيمأعلى الوسيط والنصف الآخر أسفله. الوسيط إذاً هو القيمة التي يقل عنها نصف عدد القيم ويزيد عنها النصف الآخر، أي أن 50% من القيم أقل منه و50% من القيم أعلى منه. ويعتبر الوسيط مقياساً مناسباً للنزعة المركزية في حالة البيانات الترتيبية بشكل خاص.

وللوصول إلى قيمة الوسيط، فإنه يتوجب على الباحث القيام بترتيب جميع القيم المعطاة عن المتغير إما تنازلياً أو تصاعدياً، وبعدها يتم اختيار القيمة الوسيطية التي تقسم القيم إلى مجموعتين متساويتين. وبعبارة أخرى، يمكن استخدام المعادلة البسيطة التالية في الاستدلال على موقع القيمة الوسيطية:

موقع الوسيط = $(n + 1) / 2$

حيث يشير (n) إلى عدد القيم المعطاة.

تطبق المعادلة السابقة لتحديد موقع الوسيط إذا كان حجم العينة فردياً. أما إذا كان حجم العينة زوجياً فيمكن الحصول على الوسيط من خلال حساب الوسط الحسابي للقيمتين اللتين حلتا في وسط البيانات (بعد ترتيبها تصاعدياً أو تنازلياً). وفي هذه الحالة يكون موقع القيمة الوسيطية الأولى هو $n/2$ وموقع القيمة الوسيطية الثانية هو $(n+2)/2$.

ومن أهم مزايا الوسيط أنه لا يتأثر بالقيم الشاذة أو المتطرفة بالإضافة إلى سهولة الحساب. أما أهم عيوبه فيكمن في كونه لا يأخذ عند حسابه كل القيم في الاعتبار.

2-3. المنوال ***Mode***

يعتبر المنوال ***Mode*** أبسط مقاييس النزعة المركزية. وهو يمثل القيمة الأكثر شيوعاً أو تكراراً. ولتحديد المنوال، فإن على الباحث أن يحصر التكرارات الخاصة بكل قيمة من القيم التي يأخذها المتغير المدروس. وبالتالي فإن القيمة المنوالية تكون تلك التي تتمتع بأكبر عدد من التكرارات. وعند استخدام الأعمدة البيانية أو المدرج فإن العمود الأطول في الشكل يشير إلى قيمة المنوال.

يتميز المنوال بسهولة حسابه وإمكانية استخدامه مع جميع أنواع البيانات (اسمية، ترتيبية، مستمرة) أي أن المنوال يزود الباحث بقيمة رقمية لكل من المتغيرات ذات الطبيعة النوعية والكمية في آن واحد. بالمقابل قد قد يجد الباحث نفسه أمام بيانات لا يوجد فيها أية قيم منوالية. كما قد تتضمن بعض الحالات بيانات تحتوي أكثر من قيمة من منوالية.

3- مقاييس التشتت ***Dispersion measures***

تفيد مقاييس التشتت *Dispersion Measures* في معرفة كيفية انتشار البيانات حول نقطة التركز (الوسط الحسابي). فمن الممكن أن يكون لمجموعتين من البيانات نفس الوسط الحسابي وأن تكونا مختلفتين في انتشارهما حول الوسط الحسابي. فلو تأملنا المثال التالي الذي يمثل عدد أفراد عينتين من الأسر وهي:

عدد فراد الأسرة للعينة الأولى: 4 5 6 5

عدد أفراد الأسرة للعينة الثانية: 2 6 1 11

نجد أن الوسط الحسابي لكلا العينتين متساوٍ (5) رغم التباين الواضح في عدد أفراد الأسرة في كلتا العينتين.

تستخدم مقاييس التشتت عادة مع البيانات القياسية أو المستمرة (مدرجة أو نسب). ويزودنا علم الإحصاء بمجموعة من المقاييس في هذا المجال، منها المدى والتباين والانحراف المعياري.

1.3. المدى ***Range***

يعتبر المدى ***Range*** أكثر مقاييس التشتت بدائية. فهو يمثل ببساطة الفرق بين أعلى قيمة وأقل قيمة في القيم المعطاة عن المتغير محل الدراسة.

$$Range = X_{Largest} - X_{Smallest}$$

يتميز المدى بسهولة استخدامه وفهمه، إلا أن لا يعطي فكرة واضحة عن كيفية انتشار المتغير نظراً لاعتماده على قيمتين فقط من مجموع قيم المتغير.

2.3. التباين والانحراف المعياري Variance and standard deviation

يعرف الانحراف عن الوسط الحسابي على أنه انحراف القيمة المشاهدة عن الوسط الحسابي. لكن لدى استخدام هذا المؤشر للتعرف على تشتت القيم حول الوسط الحسابي سنلاحظ أن مجموع انحرافات القيم عن وسطها الحسابي يساوي الصفر دائماً.

لحل هذه المشكلة اقترح مقياس للتشتت أطلق عليه اسم التباين ***Variance***. ويعرف التباين بأنه متوسط مربعات الانحرافات (مقدار تشتت القيم عن وسطها الحسابي).

$$S^2 = \frac{\sum(x - \bar{X})^2}{n - 1}$$

يعتبر التباين من مقاييس التشتت الشهيرة ونلاحظ من المعادلة السابقة أنه لا يمكن أن يكون سالباً. وهو يشير إلى مدى تشتت البيانات حول الوسط الحسابي، فعندما تتجمع البيانات حول الوسط الحسابي (تشتت ضعيف) يكون التباين صغيراً، وبالعكس.

لكن ونتيجة لاعتماد التباين على تربيع الانحرافات عن الوسط الحسابي فإنه لا يتماشى مع وحدات قياس المتغير محل الدراسة. لذا وللعودة إلى وحدة القياس الأصلية يتم عادةً احتساب الجذر التربيعي للتباين، ويطلق على القيمة الناتجة اسم الانحراف المعياري ***Standard Deviation***.

يعتبر الانحراف المعياري المقياس الأشهر والأكثر استخداماً للتشتت نظراً لسهولة التعامل معه رياضياً ولكونه يأخذ كافة القيم بعين الاعتبار، لكنه يتأثر بشكل واضح بالقيم الشاذة.

جدول 3.3. مقاييس النزعة المركزية والتشتت لمتغير "الألفة بالانترنت"

Statistics

الألفة بالانترنت

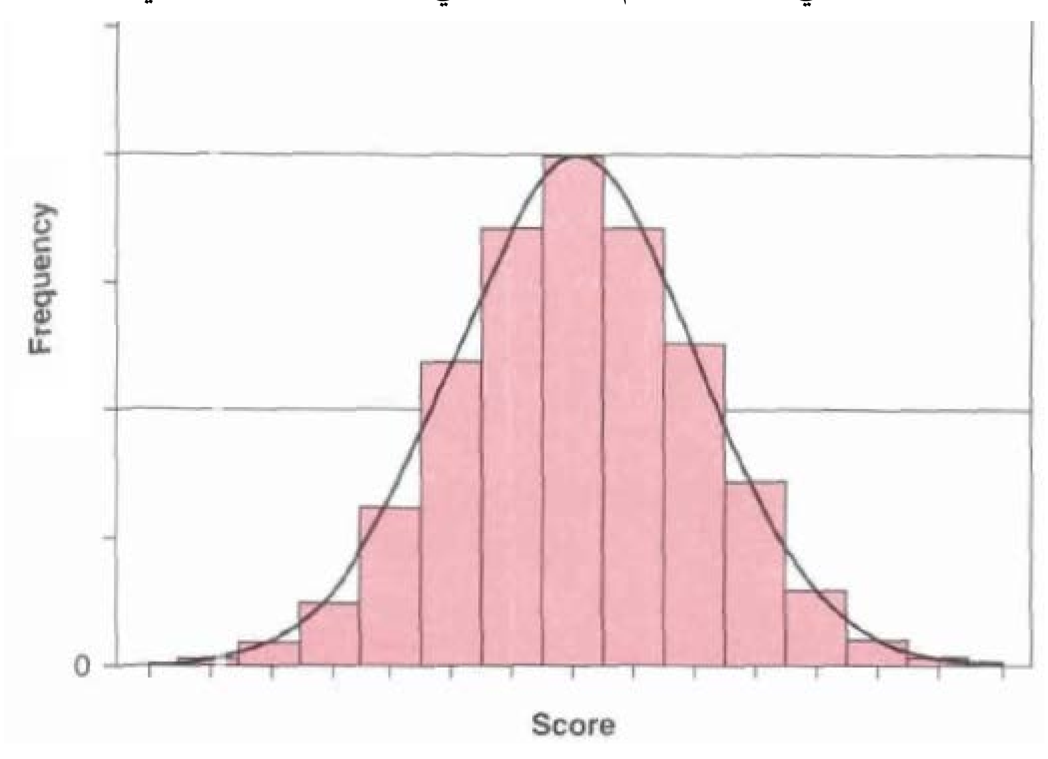
N	Valid	29
	Missing	1
Mean		4.72
Median		5.00
Mode		6
Std. Deviation		1.579
Variance		2.493
Range		5

يظهر الجدول السابق مقاييس النزعة المركزية لمتغير الألفة بالانترنت، حيث يساوي الوسط الحسابي لهذا المتغير 4.72 والوسيط 5 والمنوال 6. أما بالنسبة للتشتت فالمدى يساوي 5 والتباين 2.493 والانحراف المعياري 1.579.

4- شكل التوزيع *Distribution*

في الحالة المثالية تكون بيانات المتغير موزعة بشكل يتوافق مع شكل التوزيع الطبيعي *Normal Distribution*، حيث تتوزع البيانات بشكل متناظر تماماً حول مركزها. وعادة ما يُشبّه شكل التوزيع الطبيعي بشكل الجرس.

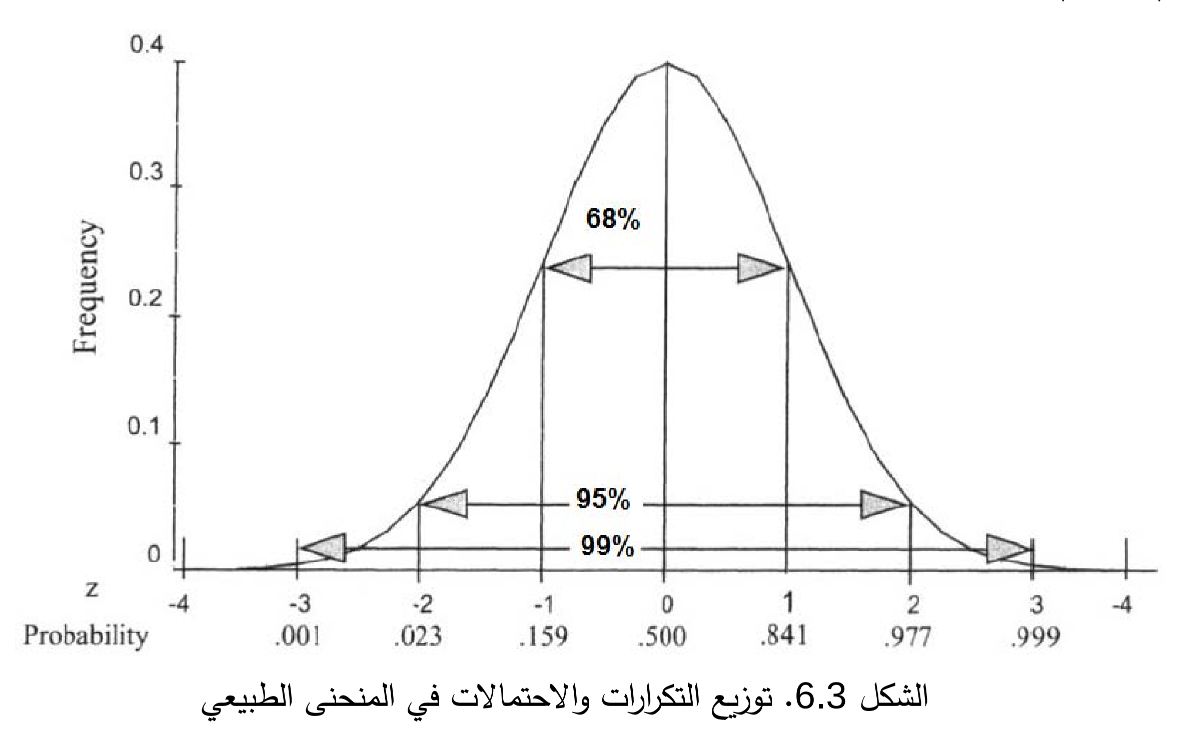
فإذا رسمنا المدرج التكراري *histogram* لبيانات موزعة بشكل طبيعي سنجد أن البيانات تتوضع بشكل متناظر حول الوسط الحسابي. كما تكون قيم الوسط الحسابي والوسيط والمنوال متساوية تماماً.



الشكل 5.3. شكل التوزيع الطبيعي

تتجمع معظم البيانات في التوزيع الطبيعي إذاً حول مركز التوزيع (أطول الأعمدة في المدرج التكراري).

وكلما ابتعدنا عن مركز التوزيع كلما قلت البيانات (كلما كان طول الأعمدة في المدرج التكراراي أصغر).
وفي حالة التوزيع الطبيعي فإن حوالي 68% من البيانات تبعد عن الوسط الحسابي بمقدار الانحراف المعيار وتبعد حوالي 95% من البيانات عن الوسط الحسابي بمقدار ضعفي الانحراف المعياري، فيما تبعد 99% من البيانات عن الوسط الحسابي بمقدار ثلاثة أضعاف الانحراف المعياري.
بمعنى آخر، ستتواجد 68% من البيانات ضمن المجال المعرف بالوسط الحسابي ± الانحراف المعياري $(\bar{X} \pm S)$. وستتواجد 95% من البيانات ضمن المجال $\bar{X} \pm 2S$.



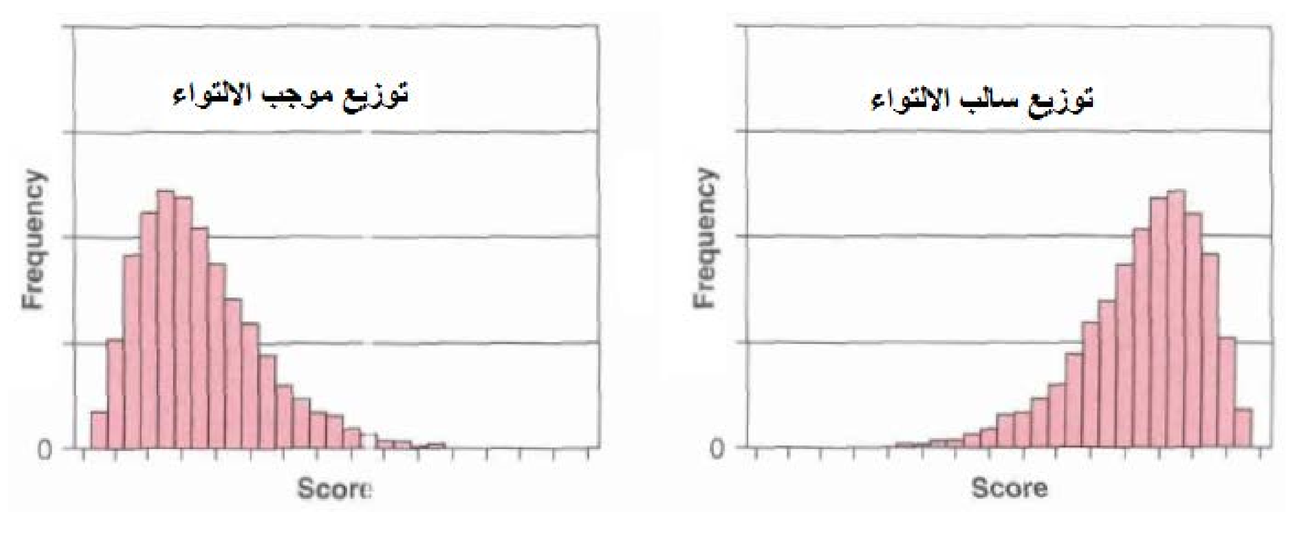
الشكل 6.3. توزيع التكرارات والاحتمالات في المنحنى الطبيعي

أما في الواقع فإن من النادر الحصول على شكل التوزيع الطبيعي المثالي، لذا وللتعرف على طبيعة توزيع البيانات (المتغير) يمكن اللجوء إلى اختبار الالتواء *Skewness* والتفلطح *Kurtosis*.

1.4. الالتواء *Skewness*

تعطي قيمة الالتواء *Skewness* فكرة عن تمركز قيم المتغير المراد معرفة شكل توزيعه. فإذا كانت قيم هذا المتغير تتمركز باتجاه القيم الصغيرة أكثر من تمركزها باتجاه القيم الكبيرة يكون شكل التوزيع لهذا المتغير ملتوٍ نحو اليمين ويسمى موجب الالتواء، وفي الحالة المعاكسة يكون التواء المتغير سالباً أو أن شكل التوزيع لهذا المتغير ملتوٍ نحو اليسار.
بمعنى آخر، عند تمثيل البيانات بالاستعانة بالمدرج التكراري *histogram* يكون التوزيع ملتوياً نحو

اليمين إذا امتد شكل التوزع أو منحنى التوزع نحو اليمين بشكل أكبر من امتداده نحو اليسار (والعكس بالعكس).



الشكل 7.3. شكل التوزيع الملتوي

وعندما يكون التوزيع ملتوياً نحو اليمين، فإن القيم المتطرفة نحو اليمين تؤثر على الوسط الحسابي بسحبه نحو اليمين. وبذلك يكون الوسط الحسابي أكبر من الوسيط. أما إذا كان التوزيع ملتوياً نحو اليسار، فإن القيم المتطرفة الصغيرة تسحب الوسط الحسابي إلى اليسار، ولهذا يكون الوسط الحسابي أصغر من الوسيط. ويكون الوسط الحسابي مساوياً للوسيط وللمنوال عندما يكون التوزيع متناظراً.

تكون قيمة معامل الالتواء مساوية للصفر إذا كان التوزيع طبيعياً بشكل مثالي. وبشكل تقريبي، يعتبر اللتواء مقبولاً إذا كانت قيمة معامل الالتواء ضمن المجال $[-1,+1]$؛أي إذا كان:

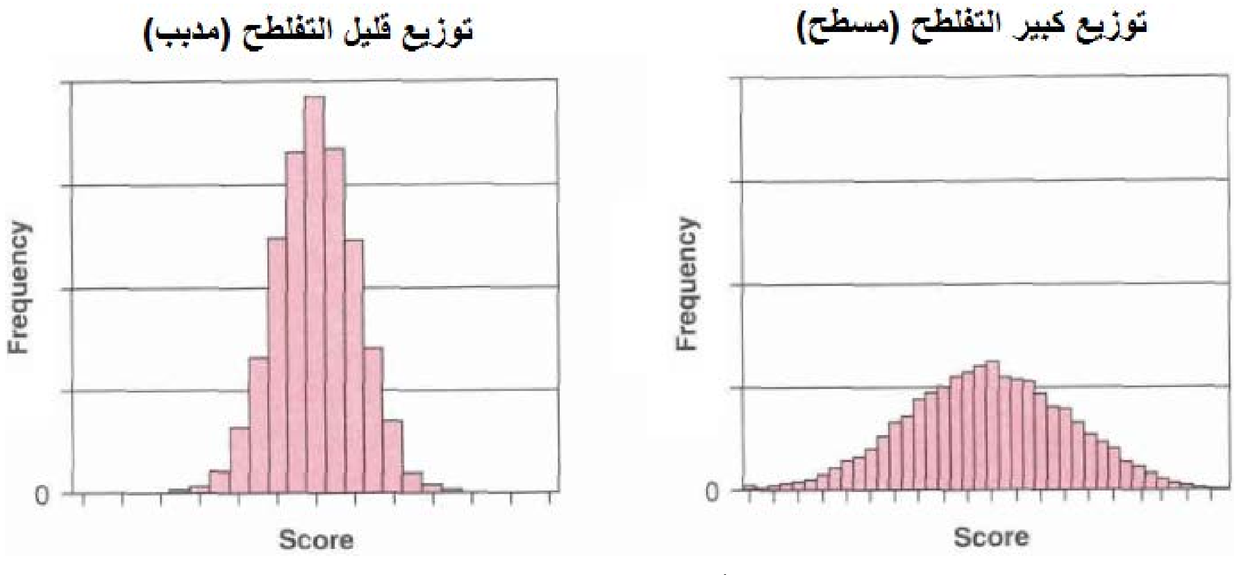
$$-1 \leq Skewness \leq +1$$

2.4. التفلطح *Kurtosis*

التفلطح هو توزيع يمثل تكرارات القيم على طرفي المتغير، وهو يمثل درجة ارتفاع التوزيع بالنسبة للتوزيع الطبيعي أو مدى تركز البيانات في وسط التوزيع أو على على طرفيه.

يدل معامل التفلطح *Kurtosis* على مدى تحدب أو تسطح شكل التوزيع. فإذا توزعت البيانات بشكل كبير باتجاه طرفي (ذيلي) المدرج التكراري *histogram* يميل التوزيع لأن يكون مسطحاً مقارنة بالتوزيع الطبيعي ويقال بأن الشكل كبير التفلطح ويكون معامل *Kurtosis* ذا قيمة سالبة.

أما إذا تركزت البيانات حول قمة التوزيع فإن التوزيع سيكون ذا شكل مدبب أو قليل التفلطح مقارنة مع شكل التوزيع الطبيعي، ويكون معامل *Kurtosis* ذا قيمة موجبة.



الشكل 8.3. أثر التفلطح على شكل التوزيع

ويعتبر التفلطح مقبولاً إذا تراوحت قيمة *Kurtosis* ضمن المجال $[-1, +1]$. ويلاحظ عدم وجود تأثير كبير للتفلطح على نتائج التحليلات الإحصائية لذا يتسامح الكثير من الخبراء مع قيم *Kurtosis* ويعتبرونها مقبولة ضمن المجال $[-3, +3]$. كما أن تأثير الالتواء والتفلطح على نتائج التحليلات الإحصائية يضعف بشكل كبير مع زيادة حجم العينة (200+).

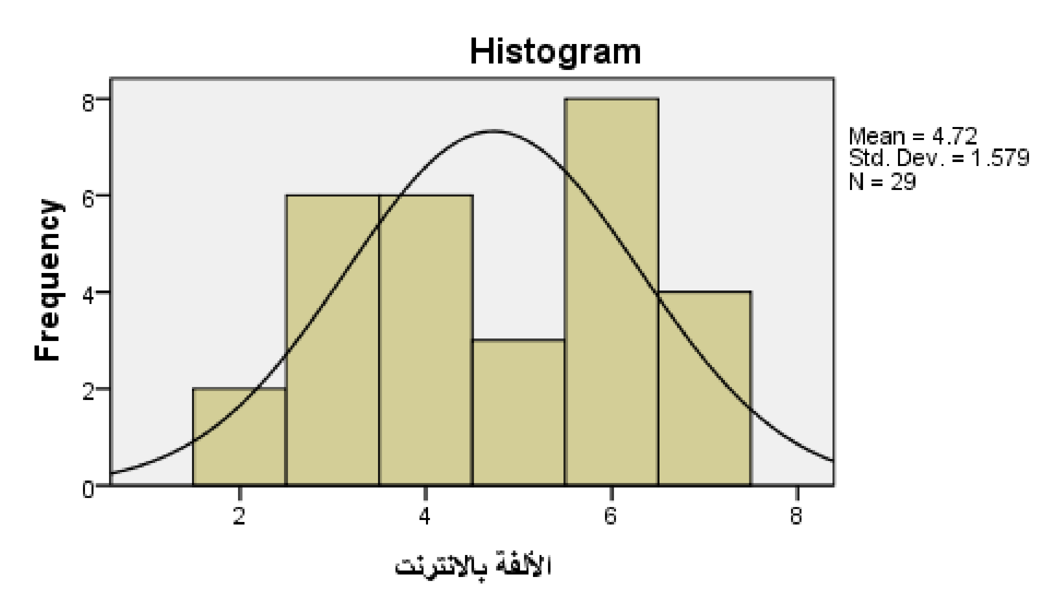
وفي *SPSS* يظهر معاملا الالتواء والتفلطح ضمن النافذة *Statistics* المنبثقة من النافذة *Frequencies*.

بالعودة إلى دراسة استخدام الانترنت للأغراض الشخصية، يظهر المدرج البياني والنتائج المتعلقة بمتغير الألفة بالانترنت أن قيمة *Skewness* تساوي -0.094 وهي تشير إلى التواء بسيط جداً نحو اليسار أي نحو القيم الصغيرة، وهذا ماتظهره العلاقة بين الوسط الحسابي والوسيط حيث أن الوسط الحسابي (4.72) أصغر من الوسيط (5) والوسيط أصغر من المنوال (6).كما يشير معامل التفلطح (*Kurtosis* = -1.261) إلى أن توزيع المتغير يميل نحو التفلطح أو التسطح نوعاً ما مقارنة بالتوزيع الطبيعي، ولكن القيمتين تبقيان ضمن الحدود المقبولة لشكل التوزيع.

Statistics

الألفة بالانترنت

N	Valid	29
	Missing	1
Mean		4.72
Median		5.00
Mode		6
Skewness		-.094-
Std. Error of Skewness		.434
Kurtosis		-1.261-
Std. Error of Kurtosis		.845



الشكل 9.3. المدرج التكراري ومعاملا الالتواء والتفلطح لمتغير "الألفة بالانترنت"

وفي النهاية لابد من الإشارة إلى وجود تحليلات إحصائية أخرى كتحليل *Kolmogrov–Smirnov* (الذي سنشرحه في فصل لاحق) لاختبار مدى التواء وتفلطح شكل التوزيع أو لاختبار التوافق بين شكل التوزيع للمتغير المدروس والتوزيع الطبيعي. إلا أن هذه التحليلات تصبح حساسة جداً إذا كبر حجم العينة (200+)، لذا ينصح *Tabachnick&Fidell (2007, p80)* باللجوء إلى اختبار شكل التوزيع من خلال المدرج التكراري في حالة العينات الكبيرة.

قائمة المراجع

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام *SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- الساعاتي، عبدالرحيم؛ حسن، أحمد السيد؛ حابس، عصام؛ البحطيطي، عبدالرحيم؛ أبو العلا، لبنى؛ الشربيني، زكريا (2009)، **تطبيقات في التحليل الإحصائي للعلوم الإدارية والإنسانية**، الطبعة الثانية، جامعة الملك عبدالعزيز، جدة.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Blumberg B., Cooper D.R., & Schindler P.S. (2005), **Business Research Methods**, Mcgraw-Hill, Berkshire.*
- *Coakes S.J. (2005), **SPSS for Windows: Analysis without Anguish**, John Wiley, Australia.*
- *Field A. (2006), **Discovering Statistics Using SPSS**, 2nd Edition, SAGE, England.*
- *Ho R. (2006), **Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS**, Chapman & Hall/CRC, USA.*
- *Landau S. &Everitt B.S. (2004), **A Handbook of Statistical Analysis Using SPSS**, Chapman & Hall/CRC Press, USA.*
- *Leech N.L., Barrett K.C., & Morgan G.A. (2005), **SPSS for Intermediate Statistics: Use and Interpretation**, 2nd Edition, Laerence Erlbaum Associates, Inc., USA.*
- *Mooi E. &Sarstedt M. (2011),**A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics**, Springer, Germany.*
- *Moore D.S., McCabe G.P., Alwan L.C., Craig B.A., & Duckworth W.M.*

*(2011), **The Practice of Statistics for Business and Economics**, 3rd Edition, W.H. Freeman and Company, England.*

- *Morgan G.A., Leech N.L., Gloeckner G.W., & Barrett K.C. (2004), **SPSS for Introductory Statistics: Use and Interpretation**, 2nd Edition, Lawrence Erlbaum Associates, USA.*
- *Pallant J. (2007), **SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows**, 3rd Edition, McGraw Hill, USA.*
- ***SPSS Statistics Base 17.0 User's Guide (2007**), SPSS Inc., USA.*
- *Zikmund W.G. &Babin B.J. (2010), **Essentials of Marketing Research**, 4th Edition, South-Western Cengage Learning, USA.*

مقترحات وتمارين للفصل الثالث

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1- ما هو الوسط الحسابي وكيف يحسب؟

(الحل في الفقرة: 1.2.2.)

2- ما هو الوسيط وكيف يحسب؟

(الحل في الفقرة: 2.2.2.)

3- متى يستخدم المدرج التكراري؟

(الحل في الفقرة: 1.)

4- كيف يحسب الانحراف المعيار ولماذا يستخدم؟

(الحل في الفقرة:2.2.3.)

5- لماذا يستخدم معامل *skewness* وما دلالة إشارته؟

(الحل في الفقرة:1.1.4.)

6- لماذا يستخدم معامل *kurtosis* وما دلالة إشارته؟

(الحل في الفقرة:1.1.4.)

7- قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.3(الوارد في الفصل). ثم قم بتطبيق مختلف أنواع التحليلات الوصفية التي تم تناولها من خلال هذا الفصل.

الفصل الرابع جداول التقاطع واختبار كاي مربع

Chapter 4: Cross tabulations and chi square test

الكلمات المفتاحية:

جداول التقاطع، اختبار كاي مربع، فرضية العدم، الفرضية البديلة، الدلالة الإحصائية، قوة العلاقة.

ملخص:

يوضح الفصل كيفية استكشاف إمكانية وجود علاقة بين متغيرين باستخدام جداول التقاطع. ويشرح خطوات اختبار الفرضيات والفرق بين فرضية العدم والفرضية البديلة ومفهوم الدلالة الإحصائية. يتناول الفصل أيضاً شروط وكيفية تطبيق وتفسير نتائج اختبار كاي مربع وكيفية الاستدلال على قوة العلاقة بين متغيرين اسميين.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- فهم كيفية توصيف العلاقة بين متغيرين باستخدام جداول التقاطع
- التعرف على خطوات اختبار الفرضيات
- فهم شروط اختبار كاي مربع وكيفية تطبيقه وتفسير نتائجه
- إدراك مؤشرات قوة العلاقة بين متغيرين اسميين

الفصل الرابع: جداول التقاطع واختبار كاي مربع

تفيد جداول التقاطع في وصف متغيرين من النوع الاسمي أو الترتيبي. وغالباً ما يترافق وصف المتغيرين مع افتراض ما عن إمكانية وجود علاقة بينهما. ولاختبار هذه العلاقة لابد من العودة إلى الاختبار الإحصائي المناسب وهو في هذه الحالة اختبار كاي مربع.

يقدم هذا الفصل مقدمة ضرورية لفهم معنى اختبار الفرضيات ويبين كيفية اختبار العلاقة بين متغيرين من النوع الاسمي من خلال اختبار كاي مربع.

1- جداول التقاطع *Cross tabulations*

كما ذكرنا في المقدمة تستعمل جداول التقاطع *Cross tabulations* لتوصيف العلاقة بين متغيرين أو أكثر من النوع الترتيبي أو الاسمي على وجه الخصوص ويمكن استخدامهما مع الأنواع الأخرى. وتعتبر جداول التقاطع من الطرق الشائعة لتحليل العلاقة بين البيانات الاسمية، لأنها:

- سهلة التفسير والفهم من قبل المدراء قليلي الخبرة بالإحصاء
- تؤدي سهولة التفسير إلى زيادة الارتباط بين نتائج البحث والقرارات والإجراءات الإدارية
- يمكن أن تقدم سلسلة من جداول التقاطع فهماً أعمق للظواهر المعقدة
- يمكن إجراء وبناء جداول التقاطع بسهولة.

تأخذ جداول التقاطع التي توصّف العلاقة بين متغيرين أشكالاً متعددة. وعادة ما يرمز لجداول التقاطع تبعاً لعدد الصفوف (الأسطر) والأعمدة في الجدول أي تبعاً لحالات كل من المتغيرين. فنجد جداول التقاطع أو الاقتران *table 2×2* في حالة متغيرين لكل منهما حالتان والجداول المتعددة التي تتكون من أكثر من صفين وعمودين. يشار إلى جدول التقاطع إذاً بالرمز ***n×m***؛ حيث ترمز *n* لعدد الصفوف و*m* لعدد الأعمدة.

في مثال "استخدام الانترنت للأغراض الشخصية" الذي سبق وشرحناه في الفصل السابق، تم تقسيم المجيبين إلى مجموعتين بحسب ساعات استخدامهم للانترنت:

- صغار المستخدمين (5 ساعات استخدام أو أقل أسبوعياً)
- وكبار المستخدمين (أكثر من 5 ساعات/أسبوع).

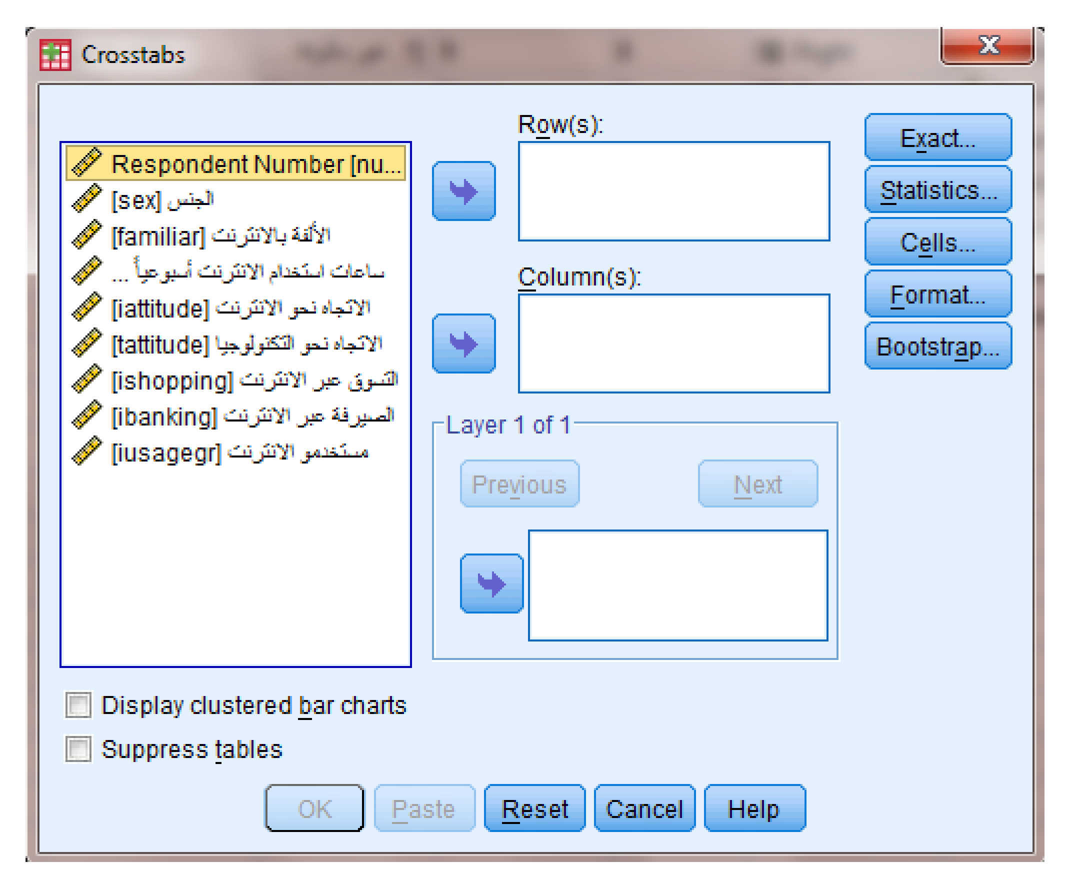
بناءً على هذا التقسيم قام الباحث بإنشاء متغير جديد باسم "مستخدمي الانترنت" وله حالتان 1="صغار المستخدمين" و2="كبار المستخدمين".

افترض الباحث وجود علاقة بين معدل استخدام الانترنت والجنس. ولاستكشاف إمكانية وجود هذه العلاقة لجأ الباحث إلى بناء جدول التقاطع الجنس × مستخدمو الانترنت.

لإنشاء جدول التقاطع في *SPSS* يمكن اتباع المسار التالي:

Analyze←Descriptive Statistics←Crosstabs

تظهر النافذة *Crosstabs*.

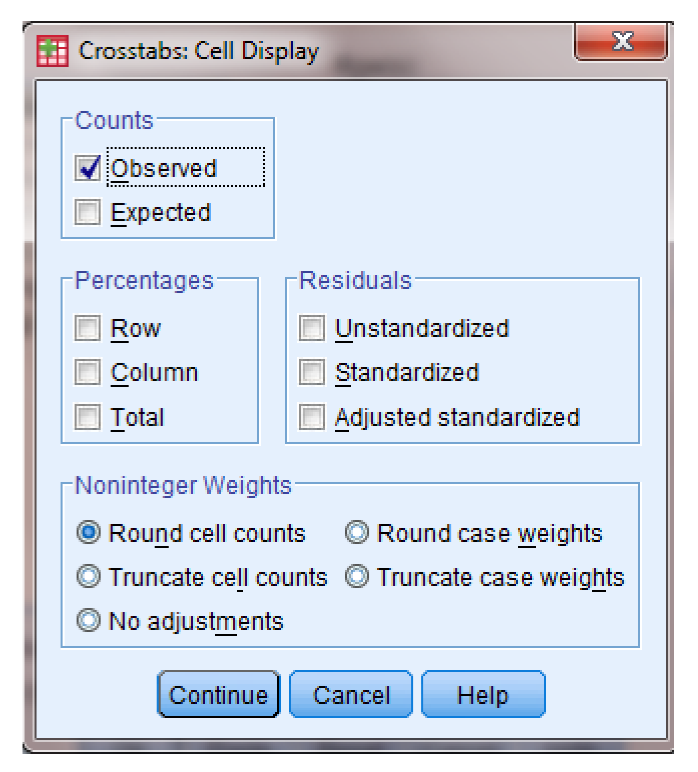


الشكل 1.4. النافذة *Crosstabs*

نسحب المتغير "الجنس" إلى *Column(s)* والمتغير "مستخدمو الانترنت" إلى *Row(s)*. وفي الحقيقة لن تختلف النتيجة إذا ماوضعنا "الجنس" في الصفوف و"مستخدمو الانترنت" في الأعمدة ولكن واصطلاحاً

ينصح بوضع المتغير المستقل في الأعمدة والمتغير التابع في الصفوف.

وإذا نقرنا فوق الزر *Cells* ستظهر النافذة التالية التي تمكننا من التحكم بطريقة عرض محتويات جدول التقاطع.



الشكل 2.4. النافذة *Cell* ضمن *Crosstabs*

تضم النافذة *Crosstabs: Cell Display* الخيارات التالية:

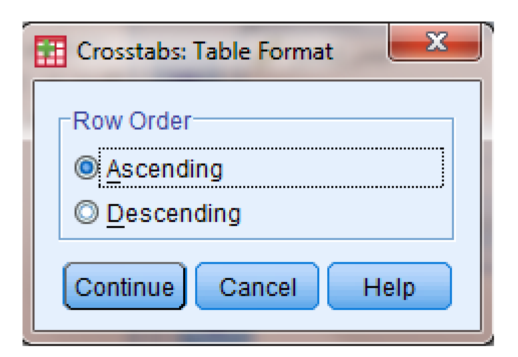
- التكرار *Counts* الذي يتضمن الخيارات التالية:
  - *Observed*: وهو الخيار الافتراضي حيث تملأ خلايا الجدول بالتكرار المشاهد O_i.
  - *Expected*: تملأ خلايا الجدول بالتكرار المتوقع E_i.
- الإطار *Percentages*: يتضمن الخيارات التالية:
  - *Rows*: تملأ خلايا الجدول بالنسب المئوية من مجموع الصف.
  - *Columns*: تملأ خلايا الجدول بالنسب المئوية من مجموع العمود.
  - *Total*: تملأ خلايا الجدول بالنسب المئوية من المجموع الكلي.
- الإطار *Residuals*:

o *Unstandardized*: تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار المتوقع $O_i E_i -$.

o *Standardized*: تملأ خلايا الجدول بالفرق بين التكرار المشاهد والتكرار المتوقع مقسوماً على الخطأ المعياري له.

o *Adjusted Standardized*: نفس الخيار السابق معبراً عنه بوحدات الانحراف المعياري عن المتوسط.

سنبقي الخيارات الافتراضية ضمن هذه النافذة حالياً. ويمكن ترتيب عرض صفوف الجدول تصاعدياً أو تنازلياً بالنقر فوق الزر *Format*.



الشكل 3.4. النافذة *Table Format*

نعود إلى النافذة *Crosstabs* وننقر فوق *OK* فتظهر النتيجة التالية.

جدول 1.4. الملخص الإحصائي

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
مستخدمو الانترنت * الجنس	30	100.0%	0	.0%	30	100.0%

يظهر الجدول الأول في النتيجة *Case Processing Summary* ملخصاً إحصائياً للمتغيرات من حيث حجم العينة والقيم المفقودة إن وجدت ونسبتها. ويظهر من الجدول أن حجم العينة هو 30 وأنه لاتوجد قيم مفقودة أو ناقصة.

جدول 2.4. جدول التقاطع الجنس × مستخدمو الانترنت

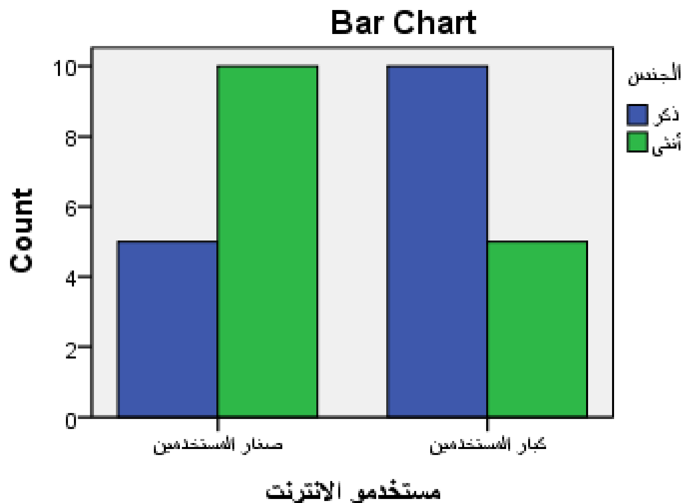
مستخدمو الانترنت * الجنس Crosstabulation

Count

		الجنس		Total
		ذكر	أنثى	
مستخدمو الانترنت	صغار المستخدمين	5	10	15
	كبار المستخدمين	10	5	15
Total		15	15	30

يبين الجدول السابق جدول تقاطع أو اقتران من النوع 2×2. يظهر الجدول مايلي:

- عدد صغار المستخدمين = 15 وعدد كبار المستخدمين =15
- عدد الذكور = 15 والإناث = 15
- عدد صغار المستخدمين (إناث)= 10 مقابل 5 ذكور
- عدد كبار المستخدمين 5 إناث مقابل 10 ذكور



الشكل 4.4. تمثيل العلاقة بين الجنس ومعدل استخدام الانترنت من خلال الأعمدة البيانية

يشير الجدول والشكل السابقان مبدئياً إلى احتمال أن يكون الذكور أكثر استخداماً للإنترنت من الإناث، لكننا لا نستطيع الاعتماد على الأرقام المشاهدة في هذا الجدول فقط للوصول إلى مثل هذه النتيجة، بل لابد من العودة إلى التحليل الإحصائي الذي يسمح لنا باختبار فرضية وجود علاقة بين الجنس ومعدل استخدام الانترنت وهو في هذه الحالة اختبار كاي مربع. ولكن وقبل تطبيق الاختبار لا بد لنا من التقديم لعملية اختبار الفرضيات.

2- اختبار الفرضيات *Hypothesis testing*

كما شرحنا في فصل سابق يستخدم الإحصاء الوصفي لوصف مجموعة من البيانات من حيث تكرار الحدوث والنزعة المركزية والتشتت، وعلى الرغم من أهمية وصف البيانات لأي تحليل إلا أن الإحصاء الوصفي لا يكفي للإجابة على الكثير من التساؤلات التي يواجهها الباحث. للإجابة على هذه التساؤلات يتوجب على الباحث الذهاب أبعد من الإحصاء الوصفي وصولاً إلى الإحصاء الاستدلالي *inferential statistics* الذي يسمح له باختبار الفرضيات. يستخدم الإحصاء الاستدلالي إذاً لتقدير قيم المجتمع الإحصائي ولاختبار الفرضيات إحصائياً.

2-1. فرضية العدم والفرضية البديلة *Null hypothesis and alternative hypothesis*

يستدعي بناء الفرضيات خضوعها للاختبار بغرض تحديد مدى صحتها، وذلك لأن البيانات تم جمعها من عينة وليس من مجتمع. والفرضية *Hypothesis* هي ادعاء حول صحة شيء ما على مستوى المجتمع. إذاً يمكن أن تكون الفرضية صحيحة أو خاطئة.

وفي اختبار الفرضيات يتم التمييز بين فرضيتين تسمى الأولى"**فرضية العدم** *Null Hypothesis*"ويرمز لها H_0 وتسمى الثانية "**الفرضية البديلة** *Alternative Hypothesis*"ويرمز لها H_1.

تعرف فرضية العدم بأنها صياغة مبدئية عن معلمة المجتمع المجهولة (وسط المجتمع مثلاً)، وهي تشير دائماً إلى عدم وجود علاقة أو اختلاف أو أثر (بحسب فرضية الباحث التي يسعى لاختبارها). تعتبر هذه الفرضية أن الاختلاف الملاحظ بين الشيئين المدروسين أو المقارنين ناتج عن الصدفة وأنه لا يوجد فرق حقيقي بينهما. وفي الحقيقية فإن فرضية العدم هي الفرضية التي يتم اختبارها ويتم رفضها عندما تتوفر دلائل على عدم صحتها.

أما الفرضية البديلة H_1 فهي الفرضية التي يضعها الباحث كبديل عن فرضية العدم، وهي تشير غالباً إلى عكس فرضية العدم أو إلى أن المعلمة المجهولة لها قيمة تختلف عن القيمة التي حددتها فرضية العدم. ويتم قبول الفرضية البديلة في حالة رفض فرضية العدم.

2-2. اختبار الفرضيات في اتجاه واحد أو اتجاهين

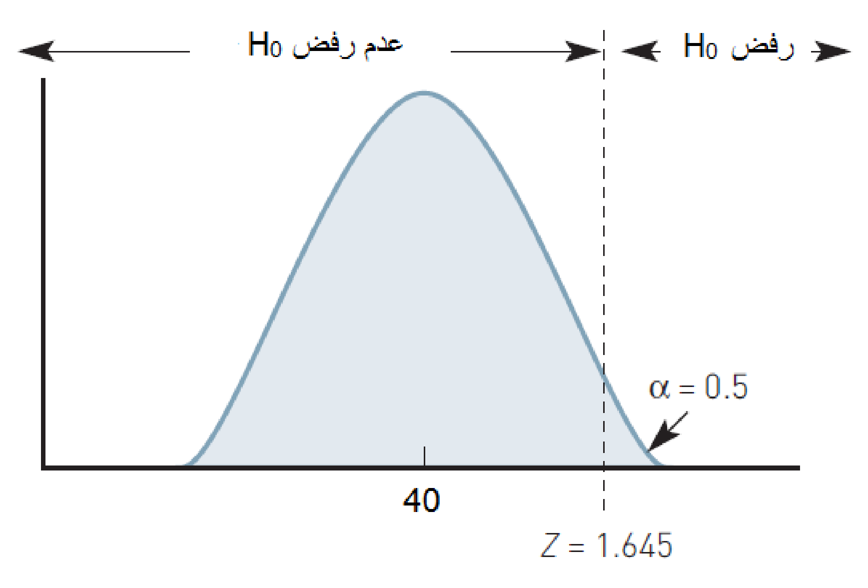
يعرف مستوى الدلالة α (ألفا) على أنه عبارة عن احتمال رفض فرضية العدم وهي صحيحة. ويعتمد تحديد مستوى الدلالة *Level of Significance* (α) على درجة استعداد الباحث لتحمل مخاطر رفض

فرضية العدم وهي صحيحة أو حجم الخطأ الذي يرضى به في قراره. وعادةً ما يتم اعتماد مستوى دلالة α مساوياً لـ 0.05 (أو 5%). وبالتالي عند اختبار الفرضيات يتم تحديد مستوى الدلالة α الذي يقيس درجة عدم التأكد، فلو كان لدينا مجال ثقة 95% يكون لدينا عدم تأكد أو خطأ مقداره 5%.

ويمكن أن يتم اختبار الفرضيات في اتجاه واحد أو اتجاهين. فاختبار الفرضيات في اتجاه واحد *One-tailed test* هو الاختبار الذي تبين فيه الفرضيات أن المعلمة للمجتمع أكبر أو أصغر من إحصائية العينة؛ فهناك إذاً تحديد للاتجاه. كمثال على ذلك، إذا أراد أحد المحلات استخدام خدمة البيع عبر الإنترنت في حال تجاوزت نسبة المتسوقين عبر الانترنت نسبة 40%، وقام بدراسة لاختبار فرضية تجاوز نسبة المتسوقين عبر الانترنت للنسبة المذكورة تكون فرضيتا العدم والبديلة في هذه الدراسة على الشكل التالي:

$$H_0 : \pi \leq 0.40$$

$$H_1 : \pi > 0.40$$



الشكل 5.4. اختبار الفرضيات في اتجاه واحد عند مستوى دلالة 5%

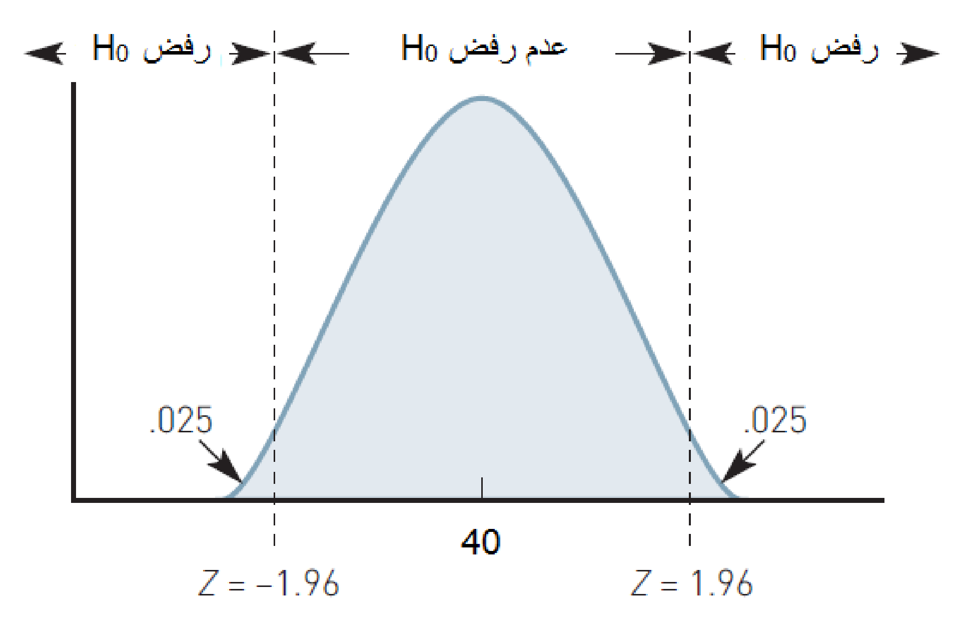
إذا ما تم رفض فرضية العدم عندئذ سيتم قبول الفرضية البديلة وسيتم إدخال خدمة البيع عبر الانترنت.

أما اختبار الفروض في جانبين فهو الاختبار الذي تبين فيه الفرضية البديلة أن المعلمة للمجتمع لا تساوي إحصائية العينة؛ أي أنه لا يوجد تحديد للاتجاه. ففي المثال السابق، إذا ما أراد المحل اختبار إذا ما كانت نسبة المتسوقين عبر الانترنت لا تساوي 40% تصبح فرضية العدم والفرضية البديلة على الشكل التالي:

$$H_0 : \pi = 0.40$$

$$H_1 : \pi \neq 0.40$$

وإذا ما تم رفض فرضية العدم عندئذ سيتم قبول الفرضية البديلة وسيتم إدخال خدمة البيع عبر الانترنت.



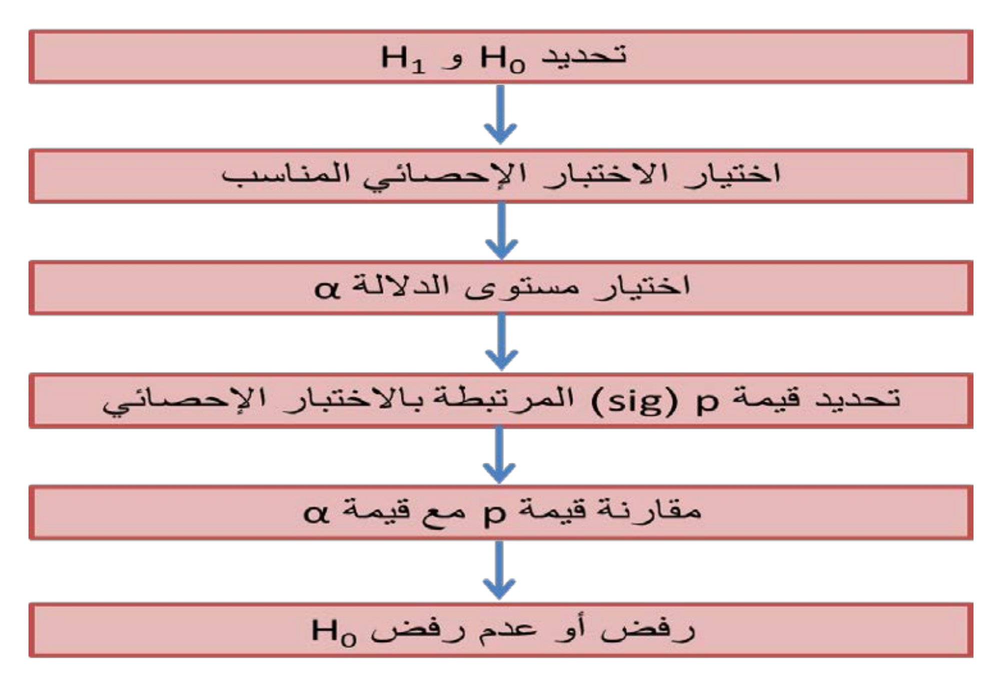
الشكل 6.4. اختبار الفرضيات في اتجاهين عند مستوى دلالة 5%

3-2. خطوات اختبار الفرضيات ***Steps of Hypothesis testing***

عند اللجوء إلى الإحصاء الاستدلالي لإجراء اختبار الفرضيات يتم بداية تحديد فرضية العدم والفرضية البديلة للاختبار. يتم تالياً جمع البيانات وتحديد الاختبار الإحصائي المناسب لا ختبار الفرضيات. يحدد الباحث أيضاً مستوى الدلالة α (عادةً 5%) ثم يتم إيجاد القيمة الاحتمالية *p value* بالاستعانة ببرامج التحليل الإحصائي (*SPSS* في هذا المقرر) حيث تشير قيمة *p* إلى مستوى الدلالة المحسوب *Computed significance level*. يقوم الباحث بعدها بمقارنة قيمة *p value* (قيمة ***sig*** ضمن *SPSS*) مع مستوى الدلالة $\alpha = 0.05$. ينتج عن هذه المقارنة الاحتمالان التاليان:

- إذا كانت قيمة *p* أصغر من α ($p < 0.05$) نرفض فرضية العدم ونقبل الفرضية البديلة.

• إذا كانت $p \geq 0.05$ ، لا يمكننا رفض فرضية العدم. والتعبير المستخدم هنا هو "**لا يمكن رفض فرضية العدم**" ولا يجوز استخدام التعبير "**تقبل فرضية العدم**"، وذلك لأنه لا يمكن أبداً إثبات فرضية العدم وبالتالي لا يمكن قبولها.



الشكل 7.4. خطوات اختبار الفرضيات

3- اختبار كاي مربع Chi square

ترجع النشأة الأولى لاختبار كاي مربع χ^2 (تلفظ *Chi-Square*) إلى البحث الذي نشره *Karl Pearson* في أوائل القرن العشرين. ويستخدم توزيع χ^2 لاختبار الفرضيات المتعلقة بالبيانات التي تكون على شكل توزيعات تكرارية، وتعتمد جميع أشكال استخدامه على أساس مقارنة التكرارات الحقيقية مع التكرارات المتوقعة وفقاً لطبيعة التوزيع الاحتمالي للبيانات.

يقدم توزيع χ^2 وسيلة لاختبار الدلالة الإحصائية لجداول التقاطع. بمعنى آخر، يختبر χ^2 الدلالة الإحصائية للعلاقة بين متغيرين اسميين خصوصاً ويمكن استخدامه مع مستويات قياس أعلى.

بمعنى آخر، يختبر χ^2 الفرق بين التوزيع المشاهد للبيانات بين الخلايا والتوزيع المتوقع لها، وكلما كان الفرق بين التوزيع المشاهد والتوزيع المتوقع أكبر كان الاحتمال أقل بأن يعزى هذا الفرق للصدفة.

وتستخدم الصيغة التالية لحساب χ^2:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

حيث:

O_i: التكرار المشاهد في الخلية i

E_i: التكرار المتوقع في الخلية i

k: عدد الخلايا

ويمكن حساب القيم المتوقعة لكل خلية باستخدام الصيغة التالية:

$$E_{ij} = \frac{R_i C_j}{n}$$

حيث:

R_i: مجموع التكرار المشاهد في السطر i

C_j: مجموع التكرار المشاهد في العمود j

n: حجم العينة

وتحسب درجة الحرية *degrees of freedom* لـ χ^2 بالصيغة:

$$d.f. = (r - 1)(c - 1)$$

حيث يمثل r عدد الصفوف الأسطر و c عدد الأعمدة.

ولكي يكون اختبار χ^2 مناسباً يجب أن نتأكد من وجود عدد كاف من المشاهدات ضمن كل خلية. وبشكل خاص، عند درجة حرية تساوي 1 يجب ألا يقل حجم الخلية المتوقع عن 5. وإذا كانت درجات الحرية أكبر من الواحد ($d.f.>1$) فلا يجب استخدام χ^2 إذا كان أكثر من 20% من التكرارات المتوقعة أصغر من 5 أو حين يكون أي من التكرارات المتوقعة أصغر من 1.

ولغرض اختبار مدى استقلالية المتغيرات بعضها عن البعض تعتمد الفرضيتان التاليتان:

- فرضية العدم H_0 القائلة باستقلالية المتغيرين المدروسين عن بعضهما (أي استقلالية الصفوف عن الأعمدة) (أو $O_i = E_i$)
- الفرضية البديلة H_1 القائلة بوجود علاقة بين المتغيرين (أو $O_i \neq E_i$)

بالعودة إلى جدول التقاطع السابق (مثال استخدام الانترنت للأغراض الشخصية) للعلاقة بين الجنس

ومعدل استخدام الانترنت ستكون فرضيتا العدم والفرضية البديلة على الشكل التالي:

H_0: لا توجد علاقة بين الجنس ومعدل استخدام الانترنت (بمعنى أن المتغيرين مستقلان وأن الاختلاف في التكرارات الظاهرة في جدول التقاطع ناتج عن الصدفة؛ أي أن معدل استخدام الانترنت لدى الذكور لا يختلف عن معدل استخدام الانترنت لدى الإناث)

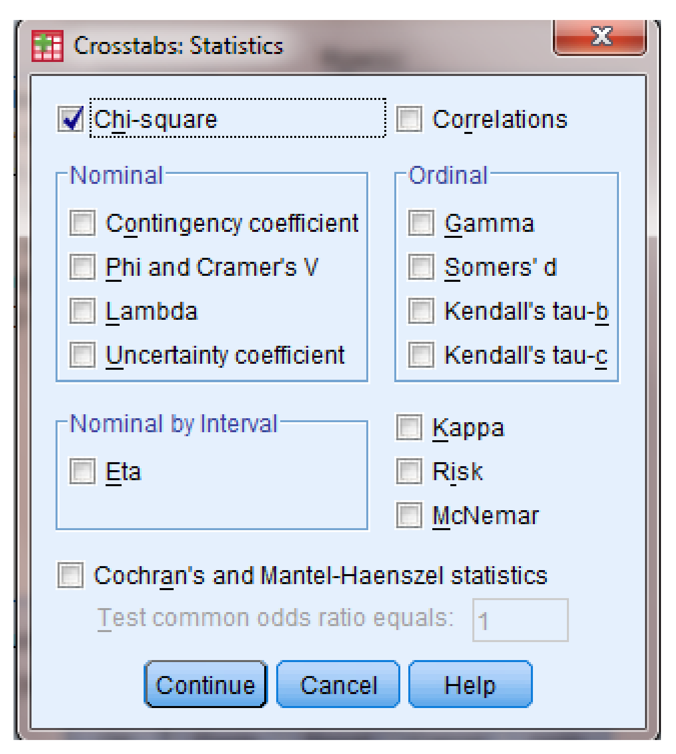
H_1: توجد علاقة بين الجنس ومعدل استخدام الانترنت

1.3. التطبيق في SPSS: Applying in SPSS

لتطبيق اختبار χ^2 في *SPSS* نتبع المسار التالي:

Analyze←Descriptive Statistics←Crosstabs←Statistics

يتم اختيار *chi–square* من النافذة *Crosstabs: Statistics*.



الشكل 8.4. النافذة *Crosstabs: Statistics*

فتظهر نتيجة اختبار χ^2 في الجدول التالي:

جدول 3.4. نتيجة اختبار χ^2

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.333[a]	1	.068		
Continuity Correction[b]	2.133	1	.144		
Likelihood Ratio	3.398	1	.065		
Fisher's Exact Test				.143	.072
Linear-by-Linear Association	3.222	1	.073		
N of Valid Cases	30				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.50.

b. Computed only for a 2x2 table

من الجدول السابق نلاحظ أن χ^2 *(1) = 3.333* وأن مستوى الدلالة المحسوب *p* (أو *sig*) $Sig = 0.068 > \alpha$؛ وبالتالي فإن فرضية العدم القائلة بعدم وجود علاقة بين الجنس ومستوى استخدام الانترنت **لا يمكن رفضها**، حيث أن هذه العلاقة ليست ذات دلالة عند مستوى دلالة 5%.بمعنى آخر يشير الاختبار إلى عدم وجود فرق بين الذكور والإناث من حيث مستوى استخدام الانترنت.

لاحظ أيضاً أن النتيجة تذكرنا (أسفل الجدول) بشرط أن ألا يقل عدد التكرارات المتوقعة ضمن خلايا جدول التقاطع عن 5، وهذا الشرط محقق حيث أن أقل تكرار متوقع هو 7.50 في مثالنا.

2.3. قوة العلاقة *Strength of association*

عندما يظهر χ^2 وجود علاقة بين المتغيرين المدروسين (أي إذا كانت $p < \alpha$) فيمكن تحديد أو قياس قوة العلاقة من خلال عدة معاملات متاحة ضمن *SPSS* ضمن النافذة *Crosstabs: Statistics*.

1.2.3. معامل الاقتران *Phi Coefficient*

يستخدم معامل الاقتران Ø (*Phi Coefficient*) لقياس قوة العلاقة أو الاقتران بين المتغيرين المدروسين في حالة الجداول المكونة من سطرين وعمودين فقط *(table 2x2)*.

تتراوح قيمة هذا المعامل بين القيم **0** والقيمة **1**. وعندما تكون قيمة المعامل 0 تكون قيمة كاي مربع مساوية للصفر أيضاً مما يدل على عدم وجود أي ارتباط بين المتغيرين المدروسين. وبالمقابل عندما يكون الارتباط تاماً بين المتغيرين المدروسين، تكون قيمة معامل الاقتران Φ مساوية 1. وعندما لا يكون هناك ارتباط بين المتغيرين بحسب اختبار χ^2 (كما هو الحال في مثالنا) لا تكون هناك حاجة لحساب

معامل الاقتران.

2.2.3. معامل التوافق *Contingency Coefficient*

يستخدم معامل التوافق *Contingency Coefficient* لقياس قوة العلاقة في جداول أكبر من *2x2* وتتراوح قيمة هذا المعامل بين القيم 0 والقيمة 1.

3.2.3. معامل كرامر *Cramer's V*

معامل كرامر *Cramer's V* هو معامل معدل من معامل الاقتران Φ ويستخدم مع جداول أكبر من جداول *2 x 2*. تتراوح قيمة هذا المعامل أيضاً بين 0 و1. وينصح باستخدامه بشكل أكبر من معامل التوافق.

أخيراً، لا بد من الإشارة إلى أن المعاملات السابقة تستخدم لقياس قوة العلاقة بين متغيرين اسميين أما في حال العلاقة بين متغيرين ترتيبيين فيتم اللجوء إلى معاملات أخرى مثل *Kenadall's Tau-b* وغيره، وسيتم التطرق لبعض هذه المعاملات لاحقاً في هذا المقرر.

قائمة المراجع

المراجع العربية:

- الساعاتي، عبدالرحيم؛ حسن، أحمد السيد؛ حابس، عصام؛ البحطيطي، عبدالرحيم؛ أبو العلا، لبنى؛ الشربيني، زكريا (2009)، **تطبيقات في التحليل الإحصائي للعلوم الإدارية والإنسانية،** الطبعة الثانية، جامعة الملك عبدالعزيز، جدة.
- الطويل، ليلى (2014)، **منهجية البحث العلمي،** كلية الاقتصاد جامعة تشرين، سورية.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS*،** الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Blumberg B., Cooper D.R., & Schindler P.S. (2005),* ***Business Research Methods****, Mcgraw-Hill, Berkshire.*
- *Coakes S.J. (2005),* ***SPSS for Windows: Analysis without Anguish****, John Wiley, Australia.*
- *Field A. (2006),* ***Discovering Statistics Using SPSS****, 2nd Edition, SAGE, England.*
- *Ho R. (2006),* ***Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS****, Chapman & Hall/CRC, USA.*
- *Landau S. &Everitt B.S. (2004),* ***A Handbook of Statistical Analysis Using SPSS****, Chapman & Hall/CRC Press, USA.*
- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*
- *Morgan G.A., Leech N.L., Gloeckner G.W., & Barrett K.C. (2004),* ***SPSS for Introductory Statistics: Use and Interpretation****, 2nd Edition, Lawrence Erlbaum Associates, USA.*
- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows****, 3rd Edition, McGraw Hill, USA.*
- *Zikmund W.G. &Babin B.J. (2010),* ***Essentials of Marketing Research****, 4th Edition, South-Western Cengage Learning, USA.*

<u>مقترحات وتمارين للفصل الرابع</u>

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1- لماذا تستخدم جداول التقاطع؟

(الحل في الفقرة: 1.)

2- ما هي خطوات اختبار الفرضيات؟

(الحل في الفقرة: 3.2.)

3- ما هي فرضية العدم والفرضية البديلة في اختبار كاي مربع؟

(الحل في الفقرة: 3.)

4- متى يستخدم معامل الاقتران وما هي دلالته؟

(الحل في الفقرة:3.2.1.)

5- قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.3. (الفصل الثالث) ثم قم باختبار العلاقة بين الجنس والتسوق عبر الانترنت ومن ثم اختبر العلاقة بين الجنس والصيرفة عبر الانترنت.

الفصل الخامس اختبار *t*

Chapter 5: t-test

الكلمات المفتاحية:

مقارنة المتوسطات الحسابية، اختبار *t* للعينة الواحدة، اختبار *t* للعينات المستقلة، اختبار *t* للعينات المزدوجة.

ملخص:

يشرح الفصل أنواع اختبار t وشروطه وكيفية تطبيقه ضمن *SPSS*. ويتناول اختبار *t* للعينة الواحدة لمقارنة الوسط الحسابي لمتغير كمي مع قيمة ثابتة. كما يشرح كيفية استخدام اختبار *t* للعينات المستقلة لمقارنة الوسطين الحسابيين لمتغير كمي بين مجموعتين. ويتناول الفصل أخيراً كيفية تطبيق اختبار *t* للعينات المزدوجة.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- الإلمام بأنواع اختبار *t*
- التعرف على كيفية تطبيق مختلف أنواع اختبار *t* ضمن *SPSS*
- إدراك طريقة مقارنة الوسط الحسابي لمتغير كمي مع قيمة ثابتة
- فهم كيفية مقارنة الوسطين الحسابيين لمتغيري كمي واحد بين مجموعتين
- فهم كيفية مقارنة الوسطين الحسابيين لمتغيرين كميين ضمن العينة الواحدة.

الفصل الخامس: اختبار *t*

يعتبر اختبار *t* من الاختبارات الإحصائية الشائعة والمهمة والتي تستخدم بشكل واسع من قبل الباحثين لقياس الفروقات المعنوية بين المتوسطات؛ أي أن الاختبار الإحصائي *t*يستخدم لاختبار فرضية تتعلق بالوسط الحسابي.

ومن أهم الشروط العامة الواجب توفرها لتطبيق اختبار *t*مايلي:

- يجب أن تكون البيانات من النوع المدرج *interval* أو النسب *ratio*.
- يجب أن تكون العينة عشوائية وقيم مفرداتها لا تعتمد على بعضها البعض.
- يجب أن يتبع توزيع المتغير المراد إجراء الاختبار على وسطه، التوزيع الطبيعي *Normal Distribution*.

سنتناول في هذا الفصل الأشكال الثلاثة لاختبار *t*، وهي:

- اختبار *t* للعينة الواحدة *One sample t–test*
- اختبار *t* للعينات المستقلة *Independent samples t–test*
- اختبار *t* للعينات المزدوجة *Paired sample t–test*

1- اختبار *t* للعينة الواحدة *One sample t-test*

يستخدم هذا اختبار *t* للعينة الواحدة *One Sample t–test* في الكشف عن وجود اختلاف معنوي (ذي دلالة إحصائية) للوسط الحسابي لمتغير ما لعينة واحدة عن قيمة محددة ثابتة. وعادةً ما يستخدم لاختبار الدلالة الإحصائية للفرق بين الوسط الحسابي المحسوب على مستوى العينة $\bar{X}$والوسط الحسابي المفترض للمجتمع μ.

يمكننا مثلاً اللجوء إلى اختبار *t* للعينة الواحدة للإجابة على تساؤلات من نمط:

- هل تتجاوز حصة المنتج في السوق 15%؟
- هل يتجاوز مستوى الرضا للموظفين المستوى الوسطي (على مقياس من 7 درجات)؟
- هل تتمتع العلامة التجارية بصورة إيجابية في أذهان المستهلكين (هلى يتجاوز الوسط الحسابي للصورة الذهنية وسطي المقياس أو القيمة الحيادية للمقياس)؟
- وغيرها من التساؤلات المشابهة.

1-1. الصيغة الرياضية والفرضيات ***Mathematical formula and hypothesis***

يظهر الجدول التالي كيفية صياغة فرضية العدم والفرضية البديلة المتعلقتان باختبار t للعينة الواحدة:

جدول 1.5. فرضية العدم الفرضية البديلة لاختبار t للعينة الواحدة

اختبار t في جانب واحد *One tailed t-test*	اختبار t في جانبين *Two tailed t-test*
$H_0: \mu \geq a$ أو $H_0: \mu \leq a$ $H_1: \mu > a$ $H_1: \mu < a$	$H_0: \mu = a$ $H_1: \mu \neq a$
حيث μ ترمز إلى الوسط الحسابي للمجتمع a قيمة ثابتة	

ويتم حساب قيمة t من خلال الصيغة التالية:

$$t = \frac{(\bar{X} - \mu)}{S/\sqrt{n}}$$

حيث:

$\bar{X}$: الوسط الحسابي للعينة

S: الانحراف المعياري

n: حجم العينة

وتحسب درجة الحرية $d.f.$ لاختبار t للعينة الواحدة كمايلي:

$$d.f. = n - 1$$

2-1. الاختبار باستخدام SPSS: ***Test using SPSS***

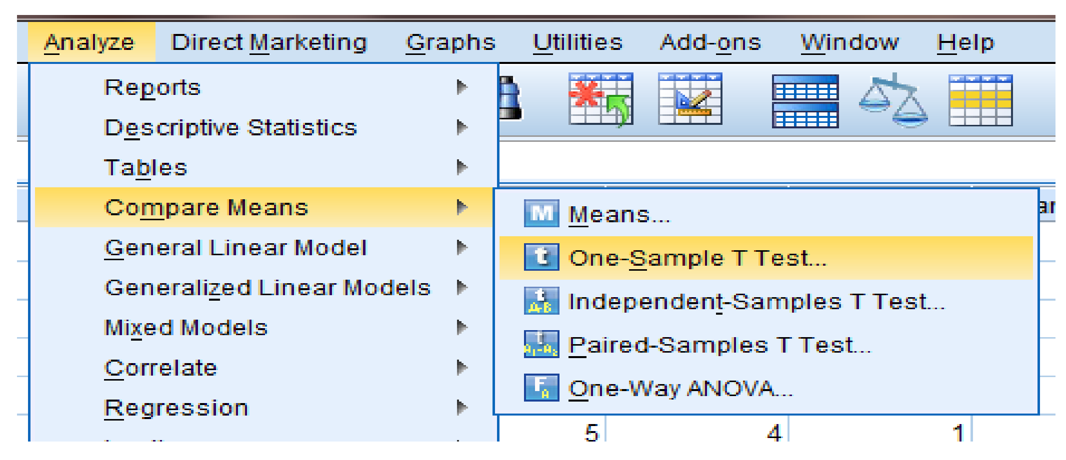
في مثال استخدام الانترنت للأغراض الشخصية يرغب الباحث في معرفة فيما إذا كان الأفراد يتمتعون بألفة جيدة بالانترنت. بمعنى آخر، هل يميل الوسط الحسابي للألفة بالانترنت نحو الجانب الموجب؟ أي هل يتجاوز هذا الوسط الحسابي وسط المقياس المستخدم لقياس الألفة بالانترنت أو القيمة 4 (نظراً لاستخدام الباحث لمقياس من 7 درجات لقياس الألفة بالانترنت)؟

فرضية العدم: $H0: \mu \leq 4$

الفرضية البديلة: $H1: \mu > 4$

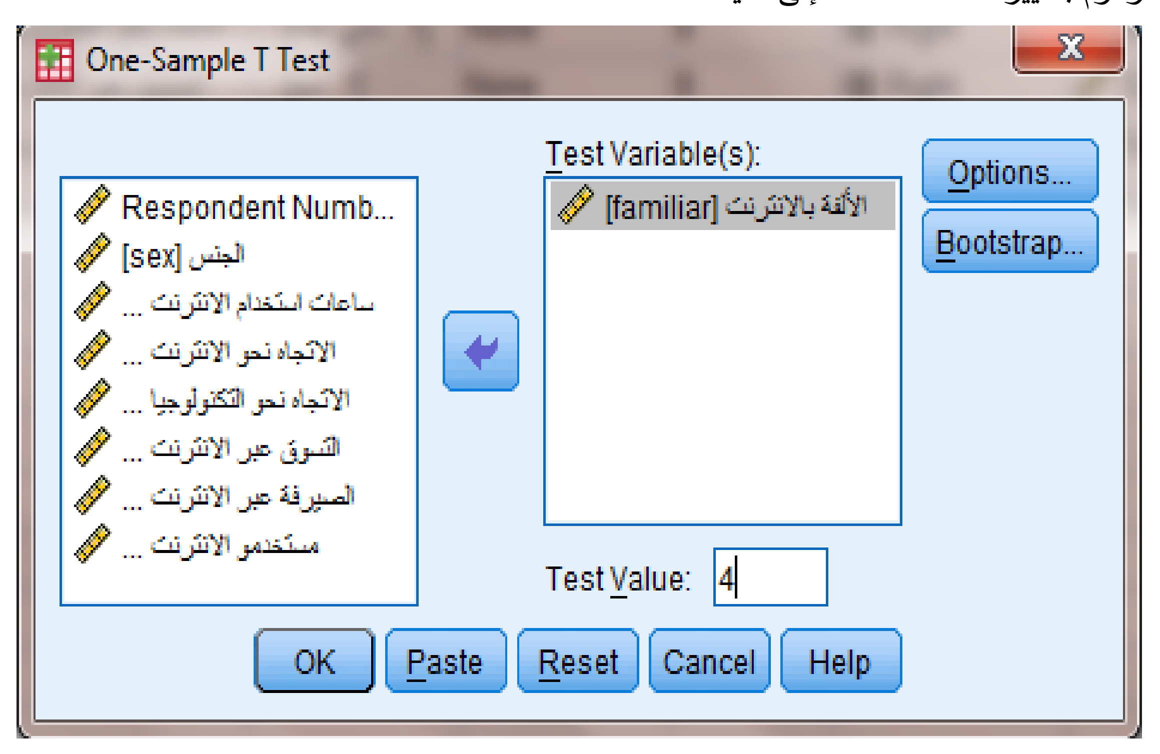
ولتنفيذ الاختبار في *SPSS* نتبع الخطوات التالية:

Analyze ←Compare Means ←One sample T Test



الشكل 2.5. خطوات تنفيذ اختبار *t* للعينة الواحدة في *SPSS*

في النافذة *One–Sample T Test* نقوم بنقل متغير "الألفة بالانترنت" إلى المربع *Test Variable(s)* ونقوم بتغيير *Test Value* إلى القيمة "**4**".



الشكل 3.5. النافذة *One-Sample T Test*

نقوم بالنقر فوق *OK* فتظهر النتيجة التالية:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
الألفة بالانترنت	29	4.72	1.579	.293

One-Sample Test

	Test Value = 4					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
الألفة بالانترنت	2.470	28	.020	.724	.12	1.32

الشكل 4.5. نتيجة اختبار *t* للعينة الواحدة

يظهر الجدول الأول *One-Sample Statistics* أن الوسط الحسابي للألفة بالانترنت يبلغ 4.72. ويؤكد اختبار *t* للعينة الواحدة في الجدول *One-Sample Test*أن الوسط الحسابي للألفة بالانترنت أكبر من وسط المقياس (القيمة 4) بشكل ذي دلالة إحصائية حيث أن $t(28) = 2.470$ و $\boldsymbol{Sig = 0.020}$ وهي أصغر من مستوى الدلالة $\boldsymbol{\alpha = 0.05}$.بمعنى آخر يمكننا أن نرفض فرضية العدم ونقبل الفرضية البديلة التي تدل على أن الأشخاص يتمتعون بمستوى إيجابي من الألفة بالانترنت.

2- اختبار *t* للعينات المستقلة *Independent samples t-test*

يستخدم اختبار *t* للعينات المستقلة *Independent Samples t-test* لفحص فرضية مساواة الوسط الحسابي لمتغير ما بين عينتين أو مجموعتين مستقلتين (الوسط الحسابيلمجموعتين). ويمكن استخدام اختبار *t* للعينات المستقلة للإجابة على تساؤلات مثل:

- هل يختلف المستخدمون عن غير المستخدمين في تقييمهم للعلامة التجارية؟
- هل ينفق ذوو الدخل المرتفع على التسلية أكثر مما ينفقه ذوو الدخل المنخفض؟
- هل يكون مستوى تحفيز الموظفين الراضين أعلى من مستوى تحفيز الموظفين غير الراضين
- هل يختلف مستوى رضا قدامى المساهمين عن مستوى رضا المساهمين الجدد؟

1.2. شروط الاختبار **_Test conditions_**

بالإضافة إلى الشروط العامة التي سبق ذكرها في مقدمة الفصل لاختبار t فإن لاختبار t للعينات المستقلة شرطان إضافيان:

- شرط استقلال المجموعتين: حيث يجب أن يظهر كل مجيب أو مشاهدة في مجموعة واحدة فقط من المجموعتين.
- شرط تجانس التباين: يجب أن يكون تباين متغير الاختبار متساوياً في كلا المجموعتين. ولكن *SPSS* يتيح حلاً بديلاً كما سنرى لاحقاً في حال عدم تساوي التباين بين المجموعتين. حيث يكون للاختبار شكلان:
    - o الأول في حال افتراض أن تباين المجموعتين متساوٍ
    - o الثاني في حال افتراض أن تباين المجموعتين غير متساوٍ

ويتم اختبار شرط تجانس التباين من خلال اختبار ليفين *Levene's test*.

2.2. الصيغة الرياضية والفرضيات **_Mathematical formula and hypothesis_**

يمكننا في اختبار t للعينات المستقلة التمييز بين متغيرين. يقسم المتغير الأول العينة إلى عينتين غير متداخلتين (كمتغير الجنس مثلاً). غالباً ما يكون هذا المتغير من النوع الاسمي *Nominal Variable* ويطلق عليه في *SPSS* اسم متغير التجميع *Grouping Variabl* . أما المتغير الثاني فهو المتغير المراد اختبار تساوي الوسط الحسابي له بين المجموعتين اللتان يعرفهما المتغير الأول (متغير التجميع). يكون هذا المتغير من النوع المدرج *Interval* أو النسب *Ratio*.

وتكتب فرضية العدم والفرضية البديلة في هذا الاختبار كمايلي:

فرضية العدم: $H_0: \mu_1 = \mu_2$

الفرضية البديلة: $H_1: \mu_1 \neq \mu_2$

ويمكن حساب قيمة t بتطبيق الصيغة التالية:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث:

$\bar{X}_1$: الوسط الحسابي للمجموعة الأولى

$\bar{X}_2$: الوسط الحسابي للمجموعة الثانية

S_1^2: تباين المجموعة الأولى

S_1^2: تباين المجموعة الثانية

n_1: حجم المجموعة الأولى

n_2: حجم المجموعة الثانية

وتحسب درجة الحرية لاختبار t للعينات المستقلة من خلال الصيغة التالية:

$$d.f. = n_1 + n_2 - 2$$

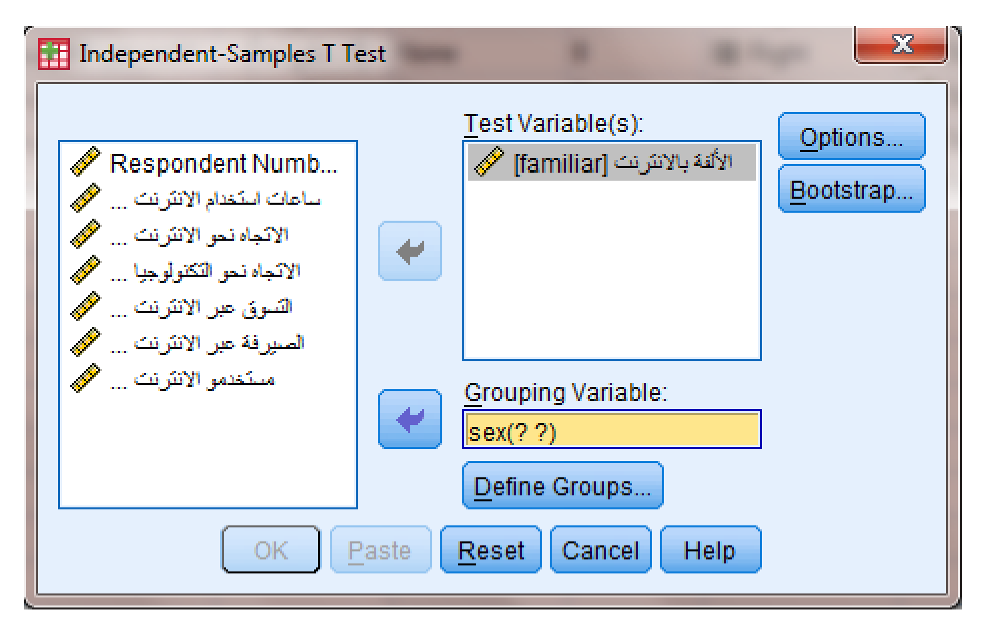
3.2. الاختبار باستخدام SPSS: Test using SPSS

في مثال استخدام الانترنت للأغراض الشخصية أراد الباحث اختبار فرضية وجود اختلاف بين الذكور والإناث من حيث الألفة بالانترنت.

لتطبيق الاختبار في *SPSS* نتبع الخطوات التالية:

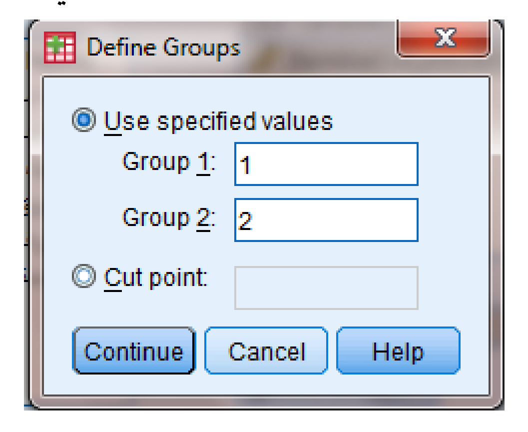
Analyze←Compare Means←Independent–Samples T Test

في النافذة *Independent–Samples T Test* نقوم بنقل المتغير الكمي "الألفة بالانترنت" إلى المربع *Test Variable(s)* وننقل المتغير الاسمي "الجنس" إلى المربع "*Grouping Variable*". نلاحظ أن الزر *Define Groups* سينشط فور القيام بنقل المتغير "الجنس" إلى المربع *Grouping Variable*.



الشكل 5.5. النافذة *Independent-Samples T Test*

ننقر فوق الزر *Define Groups* ونقوم بإدخال رمزي المجموعتين المراد مقارنة متوسطيهما (في مثالنا تم إعطاء الرمز 1 للذكور و2 للإناث) وننقر فوق الزر *Continue*. لا تظهر أهمية هذه النافذة إذا احتوى متغير التجميع على مجموعتين فقط. أما في حال احتوى متغير التجميع على أكثر من مجموعتين فتفيد هذه النافذة في تعريف المجموعتين المراد مقارنة وسطهما الحسابي.



الشكل 6.5. النافذة *Define Groups*

عند النقر فوق الزر *OK* في النافذة *Independent–Samples T Test* تظهر نتيجة اختبار *t* للعينات المستقلة.

يبين الجدول *Group Statistics* أن الوسط الحسابي لألفة الذكور بالانترنت هو 5.71 فيما يبلغ وسطي ألفة الإناث بالانترنت 3.80. ولاختبار الدلالة الإحصائية للفرق بين الذكور والإناث من حيث الألفة

بالانترنت لابد من اللجوء إلى نتيجة اختبار *t* الظاهرة في الجدول *Independent Samples Test*.

جدول 1.5. إحصاءات المجموعتين اللتان تتم مقارنتهما باستخدام اختبار *t* للعينتين المستقلتين

Group Statistics

	الجنس	N	Mean	Std. Deviation	Std. Error Mean
الألفة بالانترنت	ذكر	14	5.71	1.267	.339
	أنثى	15	3.80	1.265	.327

جدول 2.5. نتيجة اختبار *t* للعينات المستقلة

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
الألفة بالانترنت	Equal variances assumed	.015	.902	4.070	27	.000	1.914	.470	.949	2.879
	Equal variances not assumed			4.070	26.857	.000	1.914	.470	.949	2.880

نلاحظ أن الجدول يتضمن سطرين. يعطي السطر الأول *Equal variances assumed* نتيجة اختبار *t* في حال كان شرط تجانس التباين محققاً أي أن تباين المجموعة الأولى يساوي تباين المجموعة الثانية. أما السطر الثاني *Equal variances not assumed* فيظهر نتيجة اختبار *t* في حال عدم تساوي التباينين.

يسمح اختبار ليفين *Leven's Test* باختبار تساوي التباينين حيث يكون تباينا العينتين متساوياً إذا كان مستوى معنوية قيمة *F* أكبر من مستوى الدلالة المحدد ($\alpha = 0.05$) ونقوم في هذه الحالة باعتماد نتيجة اختبار *t* الموجودة في السطر الأول. أما إذا كان مستوى المعنوية (*p value* أو *Sig*) لاختبار ليفين أصغر من α فيكون تباينا المجموعتين غير متساوٍ ونلجأ لاعتماد نتيجة اختبار *t* الظاهرة في السطر الثاني.

نلاحظ من الجدول *Independent Samples Test* أن معنوية اختبار ليفين هي *sig=0.902* وهي

أكبر من مستوى الدلالة 0.05 وبالتالي فإن شرط تساوي التباينين محقق لذا نعتمد نتيجة اختبار *t* الظاهرة في السطر الأول.

تشير نتيجة الاختبار ($t(27)= 4.070$, $sig < 0.001$) إلى وجود فرق ذي دلالة إحصائية بين الذكور والإناث من حيث الألفة بالانترنت وبما أن الوسط الحسابي لألفة الذكور بالانترنت أكبر من الوسط الحسابي لألفة الإناث بالانترنت (كما يظهر في الجدول *Group Statistics*) يمكننا استنتاج أن الذكور أكثر ألفة بالانترنت من الإناث.

3- اختبار *t* للعينات المزدوجة ***Pairedsamples t-test***

يتضمن اختبار *t* للعينات المستقلة *Paired Samples t–test* فحص فرضية تتعلق بمساواة متوسط متغيرين لنفس العينة بحيث تكون مشاهدات العينة على هيئة أزواج. ويشيع استخدام هذا الاختبار في الدراسات التي تستخدم المقاييس المكررة *repeated measures* حيث يجيب الشخص مرتين على نفس المتغير كما في حال قياس قيمة متغير ما قبل وبعد تجربة أو حدث ما. فمثلاً يمكن استخدام اختبار *t* للعينات المزدوجة لاختبار معنوية الفرق بين علامات الطلاب في مقررين مختلفين أو لاختبار معنوية الفرق بين متوسط علامات الطلاب قبل دورة التقوية وبعد دورة التقوية أو لفحص اختلاف تقييم المستهلكين لعلامتين تجاريتين مختلفتين وغير ذلك من الأمثلة المشابهة.

1.3. الصيغة الرياضية والفرضيات ***Mathematical formula and hypothesis***

تكتب فرضية العدم والفرضية البديلة لاختبار *t* للعينات المزدوجة كمايلي:

فرضية العدم: $H_0: \mu_1 = \mu_2$

الفرضية البديلة: $H_1: \mu_1 \neq \mu_2$

حيث:

μ_1: الوسط الحسابي للمتغير الأول

μ_2: الوسط الحسابي للمتغير الثاني

وتحسب قيمة *t* في هذا الاختبار من خلال الصيغة التالية:

$$t = \frac{\bar{d}}{S_d/\sqrt{n}}$$

حيث:

$\bar{d}$: وسطي الفرق بين الوسطين الحسابيين للمتغيرين

S_d: الانحراف المعياري للفرق بين الوسطين الحسابيين للمتغيرين

n: حجم العينة

وتحسب درجات الحرية للاختبار كمايلي:

$$d.f. = n - 1$$

2.3. الاختبار باستخدام SPSS: Test using SPSS

في مثال استخدام الانترنت للأغراض الشخصية لنفترض أن الباحث يريد معرفة فيما إذا كان اتجاه المجيبين نحو الانترنت مختلفاً عن اتجاههم نحو التكنولوجيا.

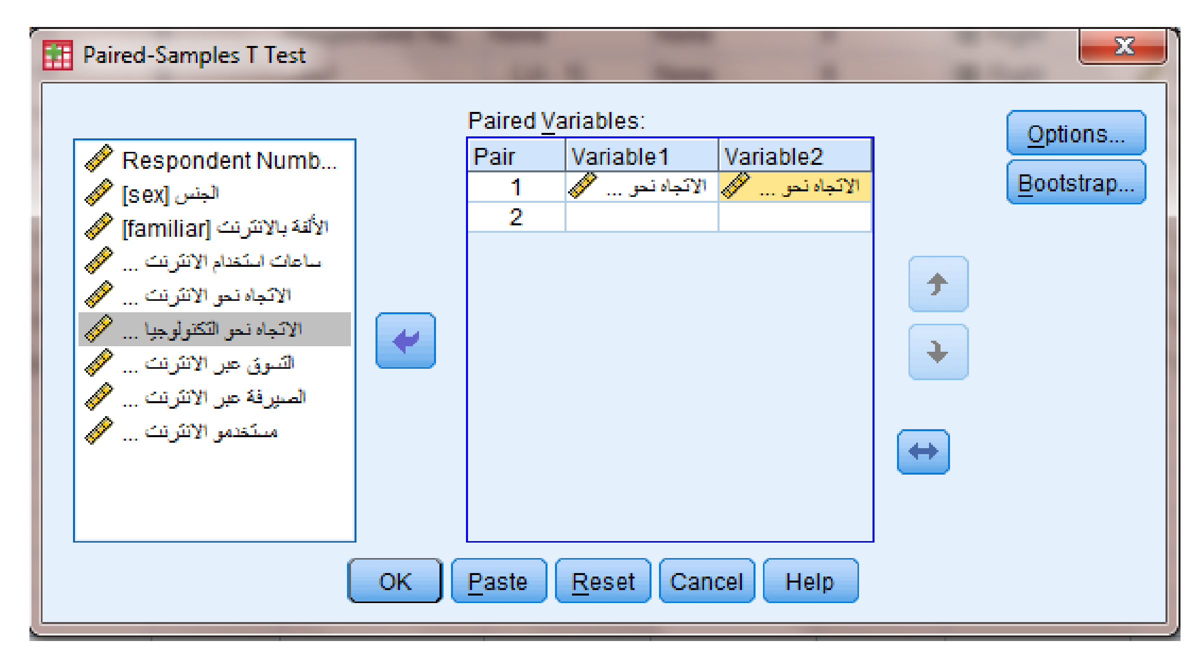
فرضية العدم: متوسط اتجاه الأفراد نحو الانترنت = متوسط اتجاه الافراد نحو التكنولوجيا

الفرضية البديلة: متوسط اتجاه الأفراد نحو الانترنت ≠متوسطاتجاه الافراد نحوالتكنولوجيا

لتطبيق الاختبار في *SPSS* نتبع الخطوات التالية:

Analyze←Compare Means←Paired-Samples T Test

في النافذة *Paired-Samples T Test* نقوم بنقل المتغير "الاتجاه نحو الانترنت" إلى العمود *Variable1* وننقل المتغير "الاتجاه نحو التكنولوجيا" إلى العمود *Variable2* ثم ننقر فوق *OK*.



الشكل 7.5. النافذة *Paired-Samples T Test*

يشير الجدول *Paired Samples Statistics* إلى أن الوسط الحسابي للعينة لمتغير "الاتجاه نحو الانترنت" هو 5.17 فيما يبلغ الوسط الحسابي لمتغير "الاتجاه نحو التكنولوجيا" 4.10 ويبلغ حجم العينة 30 مشاهدة.

جدول 3.5. الجدول الأول *Paired sample Statistics* في اختبار *t* للعينات المزدوجة

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	الاتجاه نحو الانترنت	5.17	30	1.234	.225
	الاتجاه نحو التكنولوجيا	4.10	30	1.398	.255

يشير اختبار *t* الظاهر في *Paired Samples Test* إلى وجود اختلاف جوهري بين اتجاه الأفراد نحو الانترنت واتجاههم نحو التكنولوجيا ($t(29) = 7.059$; $Sig< 0.001$). وحيث أن متوسط اتجاه الأفراد نحو الانترنت أكبر من متوسط اتجاه الافراد نحو التكنولوجيا (كما يظهر في جدول *Paired Samples Statistics*) نستطيع القول بأن اتجاه الأفراد نحو الانترنت أفضل من اتجاههم نحو التكنولوجيا. يمكن الاستفادة من هذه النتيجة وتوجيه الشركات نحو تقديم المزيد من الخدمات عبر

الانترنت حتى لو لم يكن لدى المستهلكين اتجاه إيجابي جداً نحو التكنولوجيا.

جدول 4.5. نتيجة اختبار *t* للعينات المزدوجة

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	الاتجاه نحو الانترنت - الاتجاه نحو التكنولوجيا	1.067	.828	.151	.758	1.376	7.059	29	.000

قائمة المراجع

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام** ***SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- الطويل، ليلى (2014)، **منهجية البحث العلمي**، كلية الاقتصاد جامعة تشرين، سورية.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة** ***SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Blumberg B., Cooper D.R., & Schindler P.S. (2005),* ***Business Research Methods****, Mcgraw-Hill, Berkshire.*
- *Coakes S.J. (2005),* ***SPSS for Windows: Analysis without Anguish****, John Wiley, Australia.*
- *Ho R. (2006),* ***Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS****, Chapman & Hall/CRC, USA.*
- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*
- *Malhotra N.K. &Briks D.F. (2007),* ***Marketing Research: An Applied Approach****, 3rd European Edition, Pearson Education Limited, Italy.*
- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows****, 3rd Edition, McGraw Hill, USA.*
- *Zikmund W.G. &Babin B.J. (2010),* ***Essentials of Marketing Research****, 4th Edition, South-Western Cengage Learning, USA.*

مقترحات وتمارين للفصل الخامس

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1. **ما هي الشروط الواجب توافرها لتطبيق اختبار *t*؟**

   (الحل في الفقرة: مقدمة الفصل)

2. **متى يستخدم اختبار *t* للعينة الواحدة؟**

   (الحل في الفقرة:1.)

3. **ما هي فرضية العدم والفرضية البديلة في اختبار *t* للعينة الواحدة؟**

   (الحل في الفقرة:1.1.)

4. **متى يستخدم يستخدم اختبار *t* للعينات المستقلة؟**

   (الحل في الفقرة:2.)

5. **متى يستخدم يستخدم اختبار *t* للعينات المزدوجة؟**

   (الحل في الفقرة:3.)

6. **قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.3. (الفصل الثالث) ثم قم باختبار مايلي:**

   - **هل يميل موقف الأشخاص لأن يكون إيجابياً اتجاه الانترنت؟**
   - **هل هناك فرق ذو دلالة إحصائية بين الذكور والإناث من حيث موقفهم اتجاه الانترنت؟**
   - **هل هناك فرق بين موقف الأشخاص اتجاه الانترنت وموقفهم اتجاه التكنولوجيا؟.**

الفصل السادس تحليل التباين

Chapter 6: Analysis of variance

الكلمات المفتاحية:

تحليل التباين، *ANOVA*، معامل *F*.

ملخص:

يتناول الفصل كيفية اختبار فرضية تساوي المتوسطات بين أكثر من مجموعتين من خلال اللجوء إلى تحليل التباين أو ما يعرف اختصاراً بـ *ANOVA*. ويشرح الفصل كيفية اختبار أثر المتغير المستقل (العامل) في المتغير التابع (الكمي) باستخدام *ANOVA*. كما يبين شروط تحليل *ANOVA* وفرضياته وكيفية تطبيقه ضمن *SPSS*.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- إدراك شروط وكيفية تطبيق تحليل التباين *ANOVA*.
- فهم طبيعة المتغيرات في تحليل التباين.
- التعرف على كيفية حساب المؤشرات الإحصائية ضمن *ANOVA*.
- إدراك كيفية تفسير مخرجات *ANOVA*.

الفصل السادس: تحليل التباين

استخدمنا في الفصل السابق تحليل *t* للعينات المستقلة *Independent Samples t-test* لاختبار فرضية تساوي المتوسطات بين مجموعتين. ولكن قد نحتاج في الكثير من الحالات إلى اختبار فرضية تساوي المتوسطات بين أكثر من مجموعتين (ثلاثة أو أوكثر)، يمكننا في هذه الحالة اللجوء إلى تحليل التباين *Analysis of Variance* أو ما يعرف اختصاراً بـ *ANOVA*. يعتبراً تحليل التباين إذاً امتداداً لاختبار *t* للعينات المستقلة ويمكن استخدامه لمقارنة متوسطات مجموعتين أو أكثر.

1- المتغيرات *Variables*

يسمى تحليل التباين بالأحادي (أو أحادي الاتجاه) *One-Way ANOVA* إذا كان لكل مفردة من مفردات العينة علامة أو إجابة على متغيرين. يسمى المتغير الأول المتغير العاملي أو العامل *Factor* أو المتغير المستقل *Independent Variable*. أما الثاني فهو المتغير التابع *Dependent Variable*.

1-1. المتغير المستقل *Independent variable*

يكون المتغير المستقل أو العامل عادةً متغيراً غير كمي أو غير قياسي *Non-metric Variable* أي أنه يكون متغيراً من النوع الاسمي *Nominal* أو الترتيبي *Ordinal*. يكون لهذا المتغير عدداً محدداً من الفئات أو المستويات. وهو المتغير الذي سيقسم العينة الكلية إلى عدد من المجموعات التي يراد مقارنة متوسطاتها الحسابية. ومن الأمثلة على هذا المتغير نذكر:

- مستوى التعليم: أقل من بكالوريا، بكالوريا، شهادة جامعية، ماجستير، دكتوراه.
- المحافظة: دمشق، حلب، اللاذقية،
- الحالة الاجتماعية: عازب، متزوج، مطلق.
- الجامعة: جامعة دمشق، جامعة تشرين، جامعة حلب، الجامعة الافتراضية، ...

ويمكن أن يتضمن تحليل التباين أكثر من متغير مستقل غير كمي. ففي حال كان لدينا متغير مستقل واحد نكون كما ذكرنا في حالة تحليل التباين أحادي الاتجاه، أما إذا كان لدينا متغيران مستقلان (عاملان)

فنطلق على التحليل اسم تحليل التباين ثنائي الاتجاه *Two-Way ANOVA*، وهكذا في حالة *n*متغير مستقل نكون أمام حالة *n-way ANOVA*. وسنقتصر في هذا المقرر على دراسة تحليل التباين أحادي الاتجاه.

1-2. المتغير التابع *dependent variable*

يكون المتغير التابع متغيراً من النوع الكمي أي من النوع المدرج *Interval*أو النسب *Ratio*. وهو المتغير الذي سيتم فحص مساواة متوسطه لكل فئة من فئات المتغير العاملي.

2- الفرضيات *Hypothesis*

الهدف الأساسي من تحليل التباين كما ذكرنا سابقاً هو مقارنة متوسطات متغير كمي يسمى المتغير التابع بين فئات المتغير العاملي *Factor* أو بين المجموعات التي يعرفها المتغير المستقل.
تنص فرضية العدم على أن المتوسطات التي تتم مقارنتها متساوية أي أن المتغير المستقل لا يؤثر في المتغير التابع.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_c$$

أما الفرضية البديلة فتنص على وجود اختلاف بين متوسطين على الأقل من متوسطات المجموعات التي تتم مقارنتها أي أنها تنص على وجود أثر للمتغير المستقل في المتغير.
فإذا رفضت الفرضية التي تقول إن متوسطات هذه الفئات متساوية فأي هذه المتوسطات متساوية وأيها غير متساوية؟ للإجابة على هذا السؤال تستخدم المقارنات البعدية *Post Hoc Comparisons* لمقارنة متوسطات المتغير التابع لكل زوجين من الفئات أو المجموعات على حدة.فإذا كان عدد الفئات الكلية ثلاثة، فإن عدد المقارنات البعدية سيكون ثلاث مقارنات. وبالتحديد ستكون هذه المقارنات بين المجموعتين الأولى والثانية، وبين المجموعتين الأولى والثالثة، وفي النهاية بين المجموعتين الثانية والثالثة.

3- شروط التحليل *Assumptions of the analysis*

يوجد شرطان أساسيان لاستخدام اختبار *ANOVA*:

- **التوزيع الطبيعي:** يجب أن يكون توزيع المتغير التابع (المتغير الكمي) طبيعياً لكل مجتمع من مجتمعات (مجموعات) المتغير العاملي *Factor*.وقد وجد من خلال الأبحاث أن عدم تحقق هذا الشرط لا يؤثر بشكل كبير في نتيجة تحليل التباين إذا كان عدد أفراد المجموعات التي تتم مقارنتها جيداً. وبهذا قد تكون نتيجة تحليل التباين دقيقة إلى حد ما حتى لو كان توزيع المتغير التابع غير طبيعي.
- **تجانس التباين (تساوي تباينات المتغير التابع):** يجب ألا يختلف تباين المتغير التابع بين المجموعات التي تتم مقارنتها. وكما في اختبار *t* للعينات المستقلة يتم التأكد من تحقق هذا الشرط من خلال اختبار ليفين *Leven's test*. وقد وجد أيضاً أنه يمكن الاعتماد على نتيجة *ANOVA* حتى في حال عدم تحقق هذا الشرط بشكل تام.

ولا بد من الإشارة إلى أنه ينصح بأن تكون المجموعات أو العينات التي تتم مقارنتها متقاربة من حيث عدد أفرادها.

4- حساب المؤشرات الإحصائية في *ANOVA*: *Calculating statistical coefficients in ANOVA*

تكمن الفكرة الأساسية في *ANOVA* في اختبار تغيرات المتغير التابع بين المجموعات وبناء على هذه التغيرات تحديد ما إذا كان هناك من سبب وجيه للاعتقاد بأن متوسطات مجتمعات المجموعات (أو مستويات العامل) تختلف بشكل ذي دلالة إحصائية.

يتم في *ANOVA* تجزئة التباين إلى جزئين: تباين بين المجموعات *Between-Groups Variance* وتباين ضمن المجموعات *Within-Groups Variance*. ويعتمد الاختبار على حساب المعامل *F* من الصيغة التالية:

$$F = \frac{\text{التباين بين المجموعات}}{\text{التباين ضمن المجموعات}} = \frac{Mean\ Square_{between}\ \text{متوسط مربعات الانحرافات بين المجموعات}}{Mean\ Square_{within}\ \text{متوسط مربعات الانحرافات ضمن المجموعات}}$$

$$Mean\ Square_{between} = \frac{Sum\ of\ Squares_{between}}{degrees\ of\ freadom_{between}} = \frac{\sum_{j=1}^{k} n_j(\bar{x}_j - \bar{x})^2}{k-1}$$

حيث:

$\bar{x}$: المتوسط الكلي

$\bar{x}_j$: متوسط المجموعة j

n_j: حجم المجموعة j

k : عدد المجموعات

$$Mean\ Square_{within} = \frac{Sum\ of\ Squares_{within}}{degrees\ of\ freadom_{within}} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{x}_j)}{n-k}$$

يعتمد معامل F إذاً على درجتي حرية. تتعلق الأولى بالتباين بين المجموعات $k-1$ أو عدد فئات المتغير المستقل ناقصاً واحداً. وتتعلق الأخرى بالتباين ضمن المجموعات $n-k$ أو حجم العينة ناقص عدد فئات المتغير المستقل.

إذا كانت قيمة F أصغر أو تساوي واحد فإن ذلك يدل على عدم وجود فروقات بين متوسطات المجتمعات الإحصائية وبالتالي عدم إمكانية رفض فرضية العدم. أما إذا كانت قيمة F أكبر من الواحد فلا بد من العودة إلى الدلالة الإحصائية لاختبار *ANOVA* لمعرفة إمكانية رفض فرضية العدم (p أو *sig* أصغر من α).

5- الاختبار باستخدام SPSS: *Test using SPSS*

لتوضيح كيفية تطبيق اختبار *ANOVA* أحادي الاتجاه لنأخذ المثال التالي: قامت إحدى الشركات التي تمتلك سلسلة متاجر عبر البلاد بإجراء تجربة لاختبار أثر مستوى الترويج داخل المتجر *in-store promotion* في مستوى المبيعات. تم اختبار ثلاثة مستويات من الترويج داخل المتجر:

1. مستوى ترويج مرتفع داخل المحل
2. مستوى ترويج متوسط داخل المحل
3. مستوى ترويج منخفضداخل المحل

تم اختيار 30 متجراً بشكل عشوائي لإجراء التجربة، وتم إخضاع كل 10 متاجر لإحدى الحالات الثلاثة السابقة. تم إجراء الاختبار لمدة شهرين وتم قياس المبيعات *sales* في كل متجر ومن ثم تحويل أرقام المبيعات إلى مقياس من 10 علامات. كما تم قياس معدل حضور الزبائن في المتاجر *clientele rating* على مقياس من 10 درجات. يوضح الجدول التالي البيانات التي تم جمعها للتحليل في نهاية التجربة.

جدول 1.6. بيانات تجربة أثر الترويج داخل المتجر في مستوى المبيعات

Store Number	In-Store Promotion	Sales	Clientele Rating
1	1.00	10.00	9.00
2	1.00	9.00	10.00
3	1.00	10.00	8.00
4	1.00	8.00	4.00
5	1.00	9.00	6.00
6	2.00	8.00	8.00
7	2.00	8.00	4.00
8	2.00	7.00	10.00
9	2.00	9.00	6.00
10	2.00	6.00	9.00
11	3.00	5.00	8.00
12	3.00	7.00	9.00
13	3.00	6.00	6.00
14	3.00	4.00	10.00
15	3.00	5.00	4.00
16	1.00	8.00	10.00
17	1.00	9.00	6.00
18	1.00	7.00	8.00
19	1.00	7.00	4.00
20	1.00	6.00	9.00
21	2.00	4.00	6.00
22	2.00	5.00	8.00
23	2.00	5.00	10.00
24	2.00	6.00	4.00
25	2.00	4.00	9.00
26	3.00	2.00	4.00
27	3.00	3.00	6.00
28	3.00	2.00	10.00
29	3.00	1.00	9.00
30	3.00	2.00	8.00

تسعى الشركة كما ذكرنا إلى اختبار أثر الترويج داخل المتجر في مستوى الترويج. يتمثل المتغير المستقل في هذا المثال في الترويج داخل المتجر وهو متغير اسمي ذو ثلاث فئات (مرتفع، متوسط، منخفض) أي أنه يقسم العينة إلى ثلاث مجموعات. أما المتغير التابع فهو مستوى المبيعات ضمن المحلات.

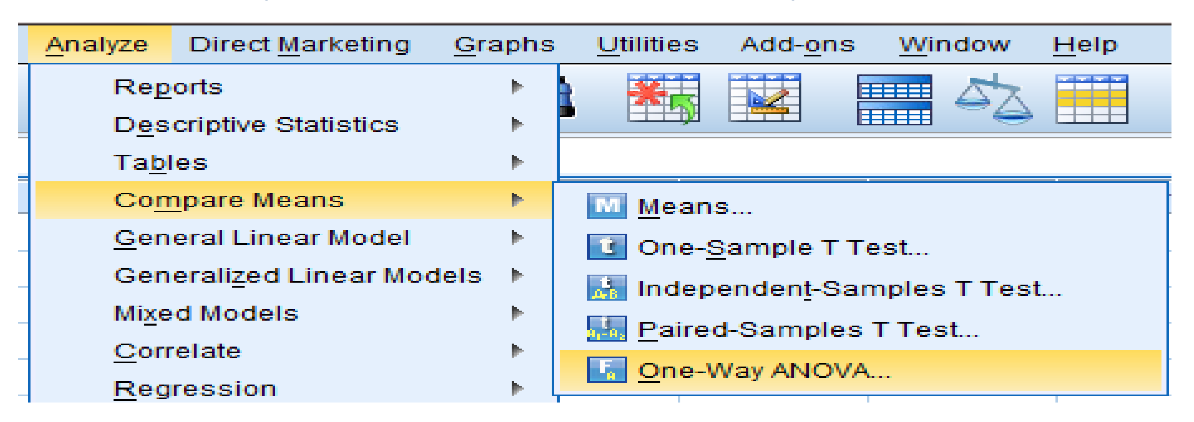
تشير فرضية العدم إلى عدم وجود أثر لمستوى الترويج داخل المتجر في مستوى المبيعات أي أنها تشير إلى عدم وجود اختلاف بين متوسطات المبيعات بين مجموعات المتاجر الثلاثة التي تم تعريفها من خلال مستوى الترويج المطبق في في المتاجر.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

أما الفرضية البديلة فتشير إلى وجود أثر لمستوى الترويج داخل المتجر في مستوى المبيعات أي أنها تشير إلى وجود اختلاف بين متوسطات المبيعات بين مستويي تروج على الأقل.

ولتطبيق الاختبار في *SPSS* يمكننا اتباع المسار التالي (كما في الشكل):

Analyze←Compare Means←One–Way ANOVA



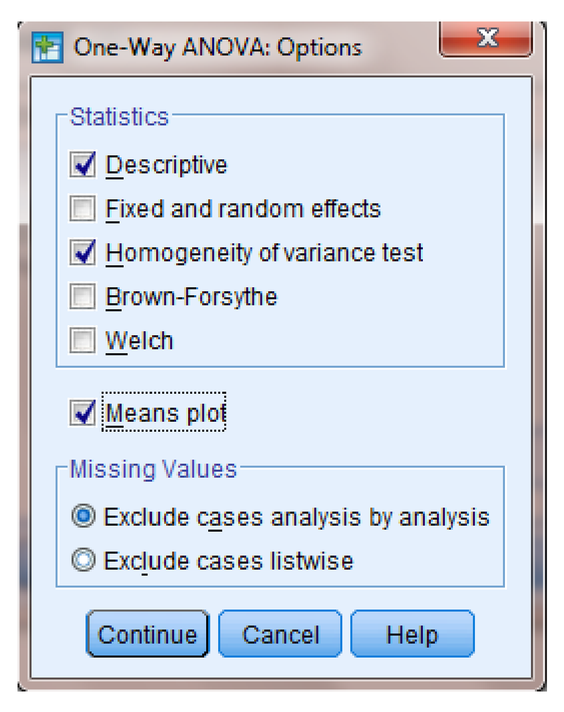
الشكل 1.6. كيفية تطبيق *ANOVA* في *SPSS*

في النافذة *One–Way ANOVA* نقوم بنقل المتغير المستقل "الترويج داخل المتجر" إلى الصندوق *Factor* وننقل المتغير التابع "المبيعات" إلى الصندوق *Dependent List*.



الشكل 2.6. النافذة *One-Way ANOVA*

وننقر فوق الزر *Options* وعند ظهور النافذة *One–Way ANOVA* نختار *Descritptive* لإظهر المتوسطات الحسابية للمبيعات في كل مجموعة ونختار *Homogeneity of Variance Test* لإظهار نتيجة اختبار تجانس التباين أو اختبار ليفين. كما يمكننا اختيار *Means Plot* للحصول على تخطيط أو تمثيل بياني لأثر الترويج داخل المتجر في مستوى المبيعات.



الشكل 3.6. النافذة *One-Way ANOVA: Options*

يبين الجدول الأول *Descriptives* في نتيجة الاختبار مجموعة من الإحصاءات الوصفية المتعلقة بمبيعات المتجر وفق مستويات الترويج الثلاثة حيث نلاحظ أن المبيعات ارتفعت بشكل ملحوظ مع ارتفاع مستوى الترويج ($M_{\text{منخفض}}=3.70 \leftarrow M_{\text{متوسط}}=6.20 \leftarrow M_{\text{مرتفع}}=8.30$) ولاختبار الدلالة الإحصائية لهذه الفروقات لا بد من العودة إلى نتيجة *ANOVA*.

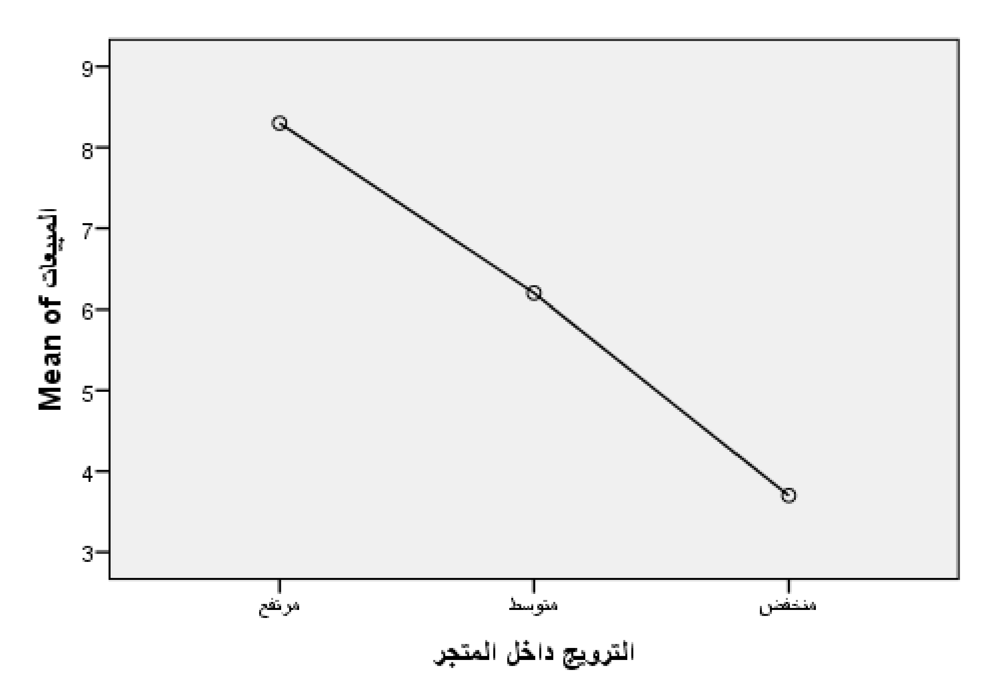
جدول 2.6. جدول الاحصاءات الوصفية *Descriptives*

Descriptives

المبيعات

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
مرتفع	10	8.30	1.337	.423	7.34	9.26	6	10
متوسط	10	6.20	1.751	.554	4.95	7.45	4	9
منخفض	10	3.70	2.003	.633	2.27	5.13	1	7
Total	30	6.07	2.532	.462	5.12	7.01	1	10

Means Plots



الشكل 4.6. الوسط الحسابي لمستوى المبيعات وفق مستوى الترويج داخل المتجر

تظهر نتيجة اختبار *Levene* في الجدول *Test of Homogeneity of Variances* أن *sig=* *0.275* وهي أكبر من)0.05(α وبالتالي فإن فرضية العدم في اختبار ليفين القائلة بتجانس أو تساوي تباين المبيعات بين المجموعات التي تتم مقارنتها لا يمكن رفضها مما يدل على تحقق شرط تجانس التباين.

جدول 3.6. اختبار تجانس التباينات

Test of Homogeneity of Variances

المبيعات

Levene Statistic	df1	df2	Sig.
1.353	2	27	.275

يظهر جدول*ANOVA* أن قيمة *P-value* (أو *sig*) المصاحبة لإحصائية *F* أقل من 0.05

F(2,27)=17,944 ; sig< 0.001

ولهذا نستطيع رفض فرضية العدم وقبول الفرضية البديلة القائلة بوجود فروق معنوية بين متوسطات

المبيعات في المتاجر تبعاً لنوع الترويج المستخدم.

جدول 4.6. نتيجة اختبار *ANOVA*

ANOVA

المبيعات

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	106.067	2	53.033	17.944	.000
Within Groups	79.800	27	2.956		
Total	185.867	29			

ولكن عند رفض فرضية العدم فإنه لا يوجد دليل واضح على وجود فروق ذات دلالة إحصائية بين كافة متوسطات المبيعات في مستويات الترويج الثلاثة فكل ما نستطيع استنتاجه من جدول *ANOVA*أن هناك فرقاً معنوياً بين متوسطي المبيعات وفق مستويي ترويج على الأقل.

ونظراً لأهمية الإجابة على هذا التساؤل وضع الإحصائيون مجموعة من الطرق التي تسمح باختبار الفروق بين متوسطات المجموعات المقارنة ومن بين هذه الطرق نجد طريقة *Bonferroni*.

حيث تسمح هذه الطريقة بإجراء مقارنات متعددة *Multiple Comparisons* لاختبار معنوية الفرق لكل زوج من فئات أو حالات المتغير المستقل (العامل).

ويمكن إظهار واختيار اختبار المقارنات البعدية بالنقر فوق الزر *Post Hoc* في النافذة *One-Way ANOVA*.

جدول 5.6. المقارنات المتعددة وفق طريقة *Bonferroni*

Post Hoc Tests

Multiple Comparisons

المبيعات
Bonferroni

(I) الترويج داخل المتجر	(J) الترويج داخل المتجر	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
مرتفع	متوسط	2.100*	.769	.033	.14	4.06
	منخفض	4.600*	.769	.000	2.64	6.56
متوسط	مرتفع	-2.100-*	.769	.033	-4.06-	-.14-
	منخفض	2.500*	.769	.009	.54	4.46
منخفض	مرتفع	-4.600-*	.769	.000	-6.56-	-2.64-
	متوسط	-2.500-*	.769	.009	-4.46-	-.54-

*. The mean difference is significant at the 0.05 level.

يظهر اختبار *Bonferroni* وجود فروق ذات دلالة إحصائية بين مبيعات المتاجر بحسب مستوى الترويج داخل المتجر فجميع المقارنات تظهر فروقاً ذات دلالة معنوية بين متوسطي المجموعتين المقارنتين. وبالعودة إلى الجدول الأول الذي يظهر متوسط المبيعات للمتاجر بحسب طريقة الترويج المتبعة نلاحظ أن المتاجر التي قامت بأعلى مستوى ترويج قد حققت أعلى مستوى مبيعات (*M=8.3*) تلتها المتاجر ذات الترويج المتوسط (*M=6.3*) فيما جاءت المتاجر التي قامت بأقل جهد ترويجي ضمن المتجر في المرتبة الأخيرة من حيث المبيعات (*M=3.7*). تدل هذه النتائج أن على المتاجر التي ترغب بزيادة مبيعاتها أن تقوم بجهود أكبر في مجال الترويج ضمن المتجر.

ولاختبار أثر الترويج داخل المتجر في معدل حضور الزبائن أعاد الباحث اختبار *ANOVA* مع استبدال المتغير التابع أي أنه استبدل المبيعات بمعدل حضور الزبائن فحصل على النتيجة التالية:

جدول6.6. اختبار أثر الترويج داخل المتجر في معدل حضور الزبائن

Descriptives

الزبائن

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
مرتفع	10	7.40	2.271	.718	5.78	9.02	4	10
متوسط	10	7.40	2.271	.718	5.78	9.02	4	10
منخفض	10	7.40	2.271	.718	5.78	9.02	4	10
Total	30	7.40	2.191	.400	6.58	8.22	4	10

Test of Homogeneity of Variances

الزبائن

Levene Statistic	df1	df2	Sig.
.000	2	27	1.000

ANOVA

الزبائن

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	.000	1.000
Within Groups	139.200	27	5.156		
Total	139.200	29			

تشير النتيجة بوضوح إلى عدم وجود أثر للترويج داخل المتجر في معدل حضور الزبائن. تعتبر هذه النتيجة منطقية على اعتبار أن الترويج داخل المتجر لن يؤثر إلا في الزبائن المتواجدين في المتجر أما خارجه فإن الزبائن لن يكونوا معرضين لهذا الترويج وبالتالي لن يؤثر هذا النوع من الترويج في اجتذاب زبائن جدد إلى المتجر.

قائمة المراجع

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام *SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- الطويل، ليلى (2014)، **منهجية البحث العلمي**، كلية الاقتصاد جامعة تشرين، سورية.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Blumberg B., Cooper D.R., & Schindler P.S. (2005),* ***Business Research Methods****, Mcgraw-Hill, Berkshire.*
- *Coakes S.J. (2005),* ***SPSS for Windows: Analysis without Anguish****, John Wiley, Australia.*
- *Ho R. (2006),* ***Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS****, Chapman & Hall/CRC, USA.*
- *Landau S. &Everitt B.S. (2004),* ***A Handbook of Statistical Analysis Using SPSS****, Chapman & Hall/CRC Press, USA.*
- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*
- *Malhotra N.K. &Briks D.F. (2007),* ***Marketing Research: An Applied Approach****, 3rd European Edition, Pearson Education Limited, Italy.*
- *Mooi E. &Sarstedt M. (2011),****A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics****, Springer, Germany.*
- *Morgan G.A., Leech N.L., Gloeckner G.W., & Barrett K.C. (2004),* ***SPSS for Introductory Statistics: Use and Interpretation****, 2nd Edition, Lawrence Erlbaum Associates, USA.*
- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data***

Analysis Using SPSS for Windows, 3rd *Edition, McGraw Hill, USA.*

مقترحات وتمارين للفصل السادس

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1- ما هي طبيعة المتغير المستقل في تحليل التباين؟

(الحل في الفقرة: 1.1.)

2- ما هي طبيعة المتغير التابع في تحليل التباين؟

(الحل في الفقرة:2.1.)

3- ما هي فرضية العدم والفرضية البديلة في تحليل التباين؟

(الحل في الفقرة:2.)

4- ما هي شروط تحليل التباين؟

(الحل في الفقرة:3.)

5- قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.6. ثم قم بإعادة اختبار أثر مستوى الترويج داخل المتجر في مستوى المبيعات.

الفصل السابع تحليل الارتباط الخطي

Chapter 7: Linear correlation analysis

الكلمات المفتاحية:

الارتباط، العلاقة الخطية، العلاقة غير الخطية، معامل بيرسون، مصفوفة الارتباط الخطي.

ملخص:

يشرح الفصل الطريقة التي يمكن من خلالها إيجاد العلاقة الخطية بين متغيرين كميين وذلك من خلال تحليل الارتباط الخطي وكيفية عرض هذه العلاقة بيانياً. ويتناول الفصل كيفية التمييز بين العلاقة الخطية والعلاقة غير الخطية بين متغيريين كميين. كما يبين كيفية استخراج مصفوفة الارتباط الخطي ضمن *spss* وشرح وتفسير مكوناتها.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- إدراك الفرق بين العلاقة الخطية والعلاقة غير الخطية بين متغيرين كميين
- إدراك شكل العلاقة الخطية بين متغيرين
- فهم كيفية اختبار العلاقة الخطية بين متغيرين كميين
- فهم مصفوفة الارتباط وتفسير محتوياتها

الفصل السابع: تحليل الارتباط الخطي

ناقشنا سابقاً اختبار أثر متغير أو أكثر ذي فئات على متغير كمي (تابع) من خلال اختبار *t* وتحليل التباين *ANOVA*. ولكن ماذا لو أردنا فحص العلاقة بين متغيرين أو أكثر من النوع الكمي؟. في هذا الفصل سيتم عرض الطريقة التي يمكن من خلالها إيجاد العلاقة الخطية بين متغيرين كميين وذلك من خلال الارتباط الخطي *Linear Correlation* وكيفية عرض هذه العلاقة بيانياً.

1- العلاقة بين متغيرين ***Relationship between two variables***

يمكن وصف العلاقة بين متغيرين من خلال:

- إمكانية وجودها
- واتجاهها
- وقوة الارتباط
- ونوع العلاقة.

لاختبار إمكانية وجود العلاقة بين المتغيرين يمكن الاعتماد على الدلالة الإحصائية لاختبار العلاقة أو الارتباط فإذا ما وجدنا دلالة إحصائية للعلاقة ($sig<\alpha$) فيمكننا القول بوجود العلاقة بين المتغيرين.

وإذا كانت العلاقة بين متغيرين موجودة فمن المهم معرفة اتجاهها، حيث يمكن أن يكون اتجاه العلاقة موجباً أو سالباً.

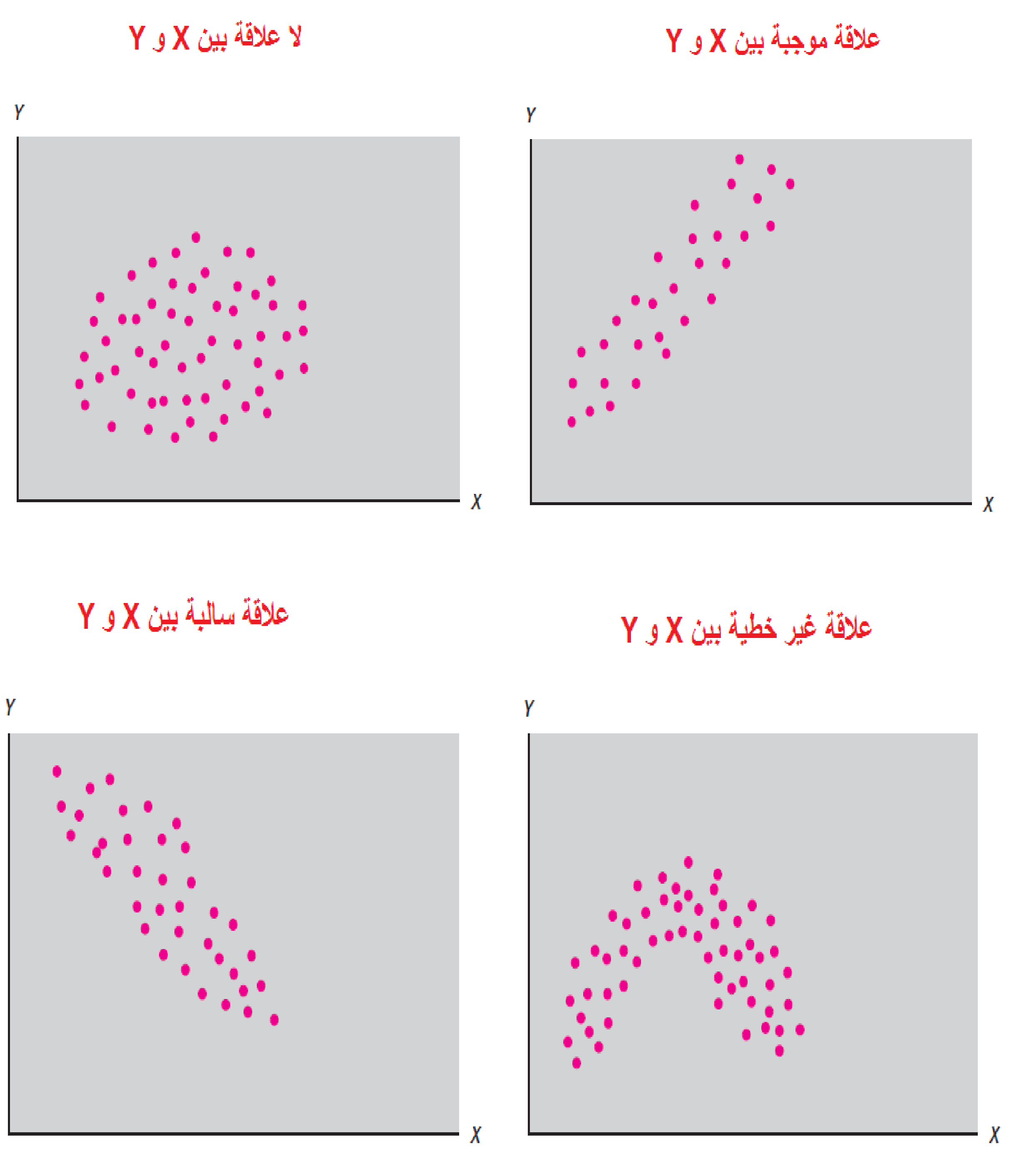
يعد فهم قوة العلاقة أمراً أساسياً أيضاً. وبشكل عام يمكن تكون العلاقة غير موجودة أو ضعيفة أو متوسطة أو قوية.

من المهم أيضاً فهم طبيعة العلاقة بين المتغيرين. إذ يمكن أن تكون العلاقة بين متغيرين X وY ذات أشكال مختلفة. فيمكن أن تكون العلاقة بين X وY علاقة خطية *Linear relationship* أي أن طبيعة وقوة العلاقة بين المتغيرين تبقى ثابتة أو متشابهة مع تطور قيم المتغيرين. ويعتبر المستقيم أفضل تمثيل للعلاقة الخطية بين متغيرين.

بالمقابل، قد تكون العلاقة بين X وY علاقة ذات شكل منحن *curvilinear relationship* (أو علاقة غير خطية) والتي تعني أن قوة و/أو اتجاه العلاقة يتغير مع تطور قيم المتغيرين. فمثلاً قد تزداد قيم Y

مع تزايد قيم X ولكن وعند نقطة معينة تبدأ قيم Y بالانخفاض مع استمرار تزايد قيم X.

ويشكل استخدام منحنى الانتشار *Scatter diagram* أحد الطرق السهلة لوصف طبيعة العلاقة بين متغيرين. يظهر الشكل أمثلة على العلاقات المحتملة بين متغيرين والتي يمكن ملاحظتها من خلال منحنيات الانتشار.



الشكل 1.7. الأشكال المحتملة للعلاقة بين متغيرين X وY

2- معامل الارتباط الخطي *Linear correlation coefficient*

يتضمن تحليل الارتباط الثنائي *Bivariate Correlation* ضمن *SPSS* ثلاثة معاملات:

- معامل ارتباط بيرسون *Pearson Correlation Coefficient*: يستخدم هذا المعامل لقياس درجة واتجاه العلاقة الخطية بين متغيرين كميين (تم قياسهما باستخدام مقياس مدرج *Interval* أو مقياس نسب *Ratio*).
- معامل ارتباط سبيرمان *Spearman Correlation Coefficient*: يستخدم لقياس درجة الارتباط (التوافق) بين متغيرين ترتيبيين *Ordinal* (تم قياسهما باستخدام مقاييس ترتيبية).
- معامل ارتباط *Kendall Tau-B*: يستخدم لقياس درجة الارتباط بين متغيرين ترتيبيين (تم قياسهما باستخدام مقاييس ترتيبية).

سيتم التركيز في هذا الدرس على معامل الارتباط الخطي بيرسون *Pearson Correlation Coefficient* والذي تم وضع صيغته الرياضية من قبل *Karl Pearson* وسمي باسمه.

2-1. شروط اختبار الارتباط الخطي *Linear correlation conditions*

عند إجراء استخدام معامل الارتباط الخطي أو معامل بيرسون لا بد من تحقق الشرطين التاليين بشكل أساسي:

- وجود علاقة خطية بين المتغيرين. وسنستعرض كيفية اختبار هذا الشرط باستخدام منحنى الانتشار في فقرة لاحقة.
- يجب أن يتمتع كلا المتغيرين المراد اختبار العلاقة بينهما بتوزيع طبيعي *Normal distribution*.

2-2. الصيغة الرياضية لمعامل بيرسون *Mathematical formula for Pearson coefficient*

يمكن حساب معامل الارتباط الخطي أو معامل بيرسون كمايلي:

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

حيث:

$\bar{X}$ الوسط الحسابي للمتغير الأول

$\bar{Y}$ الوسط الحسابي للمتغير الثاني

n حجم العينة

يعتبر معامل الارتباط الخطي أو معامل بيرسون المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية. يتراوح معامل الارتباط (r) بين -1 و$+1$. أي أن:

$$-1 \leq r \leq +1$$

تشير قيمة المعامل إلى قوة العلاقة الخطية أو قوة الارتباط الخطي بين المتغيرين أما الإشارة (− أو +) فتشير إلى اتجاه العلاقة.

في هذا السياق، لا بد من فهم المصطلحين التاليين:

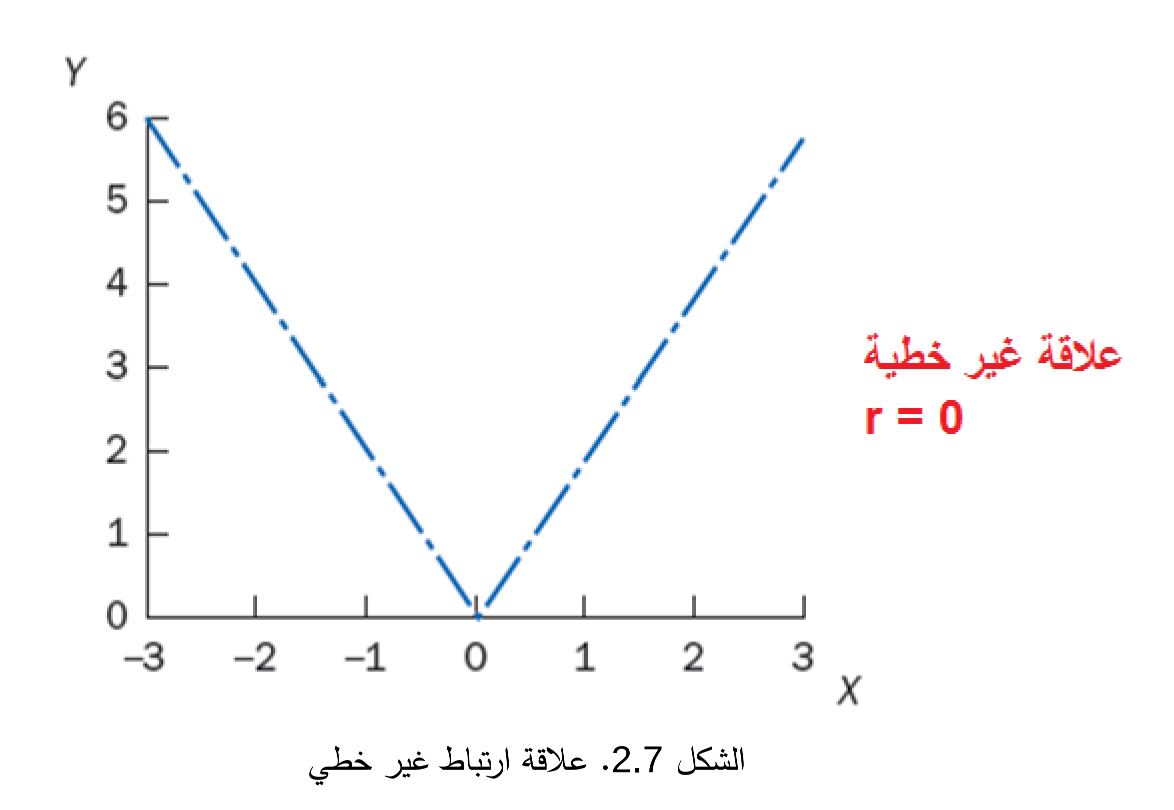
- **الارتباط الخطي الموجب:** يشير إلى أن التزايد في المتغير X يقابله تزايد في المتغير Y (أي إلى تزايد المتغيرين معاً) كما في حالة العلاقة بين السعر وجودة المنتج المدركة من قبل المستهلك.
- **الارتباط الخطي السالب:** يشير إلى أن التزايد في المتغير X يقابله تناقص في المتغير Y أي أنه يشير إلى وجود علاقة عكسية بين المتغيرين كما في حالة العلاقة بين السعر والكمية المطلوبة.

يشير معامل الارتباط $r = +1$ إلى **ارتباط خطي تام موجب** بين المتغيرين X وY أي أنه يشير إلى وجود علاقة خطية تامة بين المتغيرين X و Y. وبالتالي فإن أي تغير في X يؤدي إلى تغير بنفس النسبة ونفس الاتجاه في Y.

أما معامل الارتباط $r = -1$ فيشير إلى أن أي تغير في X يؤدي إلى تغير بنفس النسبة ولكن باتجاه معاكس في Y، ويقال أن **الارتباط بين المتغيرين هو ارتباط تام سالب**.

أما إذا كانت قيمة معامل بيرسون $r = 0$ فإن ذلك يعني **عدم وجود علاقة خطية**. بمعنى آخر، يمكن القول بأن اقتراب قيم r من ±1 يدل على قوة العلاقة الخطية بين المتغيرين. أما إذا اقتربت قيمة r من الصفر فإن ذلك يدل على ضعف العلاقة خطية بين المتغيرين. وتدل الإشارة كما ذكرنا سابقاً على اتجاه العلاقة بين المتغيرين.

نذكر هنا إلى وجوب الانتباه إلى أن الحصول على قيم صغيرة (قريبة من الصفر) لمعامل بيرسون لايعني عدم وجود علاقة بين المتغيرين X وY وإنما يعني عدم وجود علاقة خطية بينهما. فقد يكون هناك علاقة غير خطية بين المتغيرين (علاقة من الدرجة الثانية مثلاً) كما يظهر في الشكل التالي.



الشكل 2.7. علاقة ارتباط غير خطي

3.2. قوة العلاقة *Strength of the relationship*

يمكن استخدام قيمة معامل الارتباط لوصف قوة العلاقة بين المتغيرين. يتراوح معامل الارتباط كما ذكرنا سابقاً بين −1 و +1. كيف نفسر إذاً قيمة معامل الارتباط التي تتراوح بين 0 و ±1 ؟.

اختلف المؤلفون في تفسير المعامل إلا أن *Hair et al. (2003)* قد اقترحوا الجدول التالي لتفسير هذا المعامل. فبحسب المذكورين يشير معامل الارتباط الذي يتراوح بين ±0.81 و ±1 إلى ارتباط قوي جداً بين المتغيرين. على الجانب الآخر وفي حال تراوح معامل الارتباط بين ±0.01 و ±0.20 فإنه يشير إلى ارتباط ضعيف جداً بين المتغيرين وهناك احتمال قوي بأن لا يتم رفض فرضية العدم التي تشير إلى عدم وجود علاقة ارتباط خطي بين المتغيرين (إلا في حالة العينات الكبيرة).

جدول 1.7. قوة العلاقة بناء على معامل الارتباط

معامل الارتباط	قوة العلاقة
±0.81 إلى ±1.00	علاقة قوية جداً
±0.61 إلى ±0.80	علاقة قوية
±0.41 إلى ±0.60	علاقة متوسطة
±0.21 إلى ±0.40	علاقة ضعيفة
±0.01 إلى ±0.20	علاقة ضعيفة جداً

3- الاختبار باستخدام SPSS: *Test using SPSS*

قام أحد الباحثين ببناء استبيان لقياس:

- **الموقف من المدينة:** تم قياسه على مقياس من 11 نقطة (1= لا أحب المدينة أبداً، 11= أحب المدينة كثيراً).
- **مدة الإقامة في المدينة:** عدد السنوات التي قضاها المجيب في المدينة.
- **أهمية الطقس:** مقياس من 11 نقطة (1= غير مهم على الإطلاق، 11= مهم جداً).

قام الباحث بتوزيع الاستبيان على عينة من 30 شخصاً. يظهر الجدول التالي البيانات التي حصل عليها الباحث.

يريد الباحث أن يختبر إمكانية وجود علاقة أو ارتباط بين الموقف من المدينة ومدة الإقامة فيها.

فرضية العدم: لا يوجد ارتباط بين المتغيرين على مستوى المجتمع

$$H_0 : \rho = 0$$

الفرضية البديلة: يوجد ارتباط بين المتغيرين على مستوى المجتمع

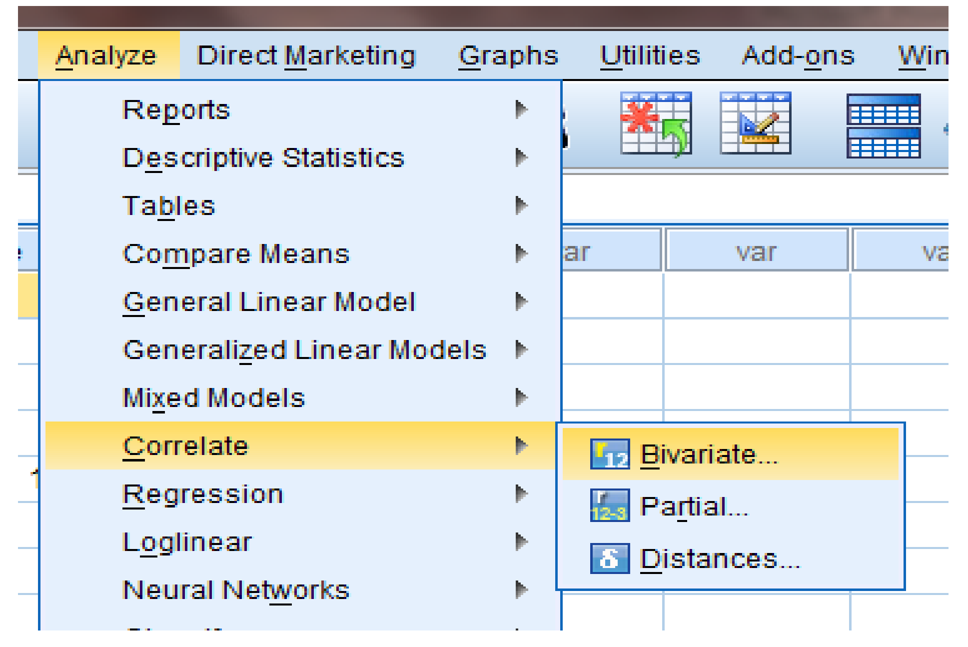
$$H_1 : \rho \neq 0$$

جدول 2.7. بيانات مثال الموقف من المدينة

المجيب	الموقف من المدينة	مدة الإقامة	أهمية الطقس
1	6	10	3
2	9	12	11
3	8	12	4
4	3	4	1
5	10	12	11
6	4	6	1
7	5	8	7
8	2	2	4
9	11	18	8
10	9	9	10
11	10	17	8
12	2	2	5
13	6	10	3
14	9	12	11
15	8	12	4
16	3	4	1
17	10	12	11
18	4	6	1
19	5	8	7
20	2	2	4
21	11	18	8
22	9	9	10
23	10	17	8
24	2	2	5
25	6	10	3
26	9	12	11
27	8	12	4
28	3	4	1
29	10	12	11
30	4	6	1

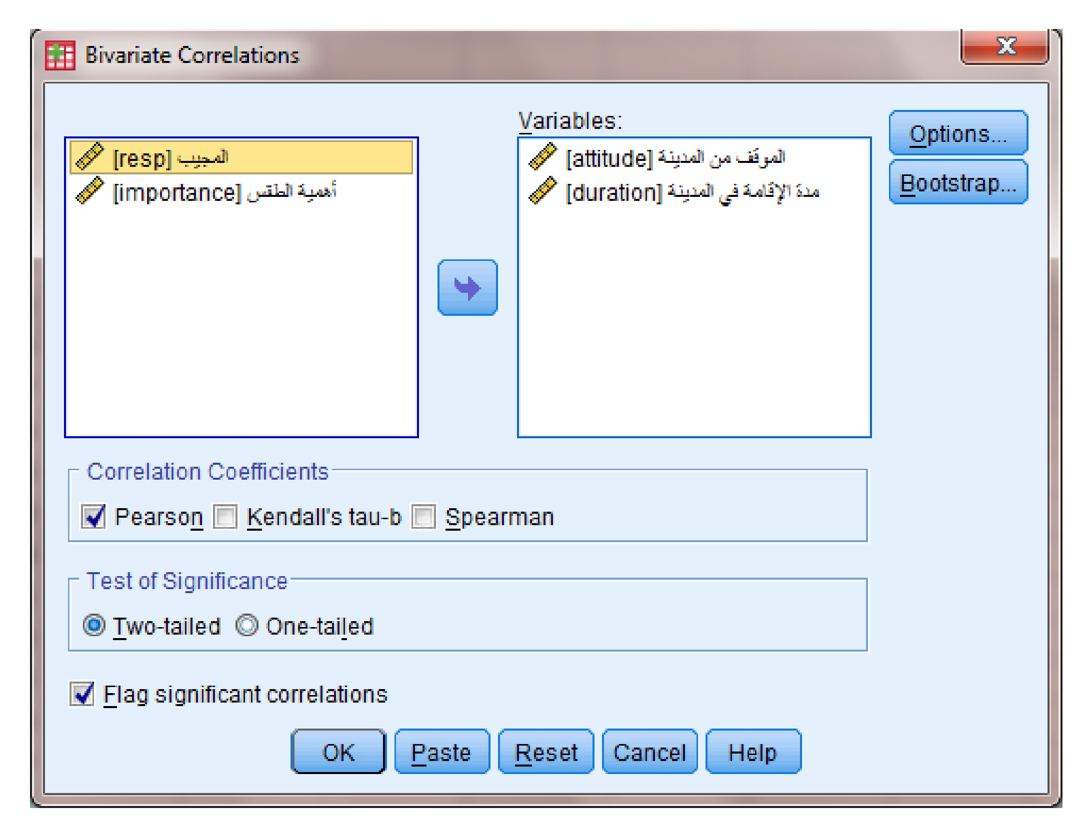
لتطبيق ختبار الارتباط الخطي في *SPSS* يمكن اتباع المسار التالي:

Analyze←Correlate←Bivariate



الشكل 3.7. كيفية تطبيق اختبار الارتباط الخطي في *SPSS*

في النافذة *Bivariate Correlations* نقوم بنقل المتغيرين اللذين نرغب باختبار الارتباط الخطي بينهما أي الموقف من المدينة *attitude* ومدة الإقامة في المدينة *duration* إلى المربع *Variables*.



الشكل 4.7. النافذة *Bivariate Correlations*

بعد النقر فوق *OK* تظهر مصفوفة الارتباط في شاشة المخرجات.

Correlations

		الموقف من المدينة	مدة الإقامة في المدينة
الموقف من المدينة	Pearson Correlation	1	.935**
	Sig. (2-tailed)		.000
	N	30	30
مدة الإقامة في المدينة	Pearson Correlation	.935**	1
	Sig. (2-tailed)	.000	
	N	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

الشكل 5.7. مصفوفة الارتباط بين الموقف من المدينة ومدة الإقامة في المدينة

نلاحظ أن قيمة معامل الارتباط الخطي *Pearson Correlation* على قطر المصفوفة تساوي 1 دائماً حيث أنه يشير إلى الارتباط بين المتغير وذاته، ومن الطبيعي أن يكون الارتباط بين المتغير وذاته ارتباطاً تاماً وموجباً. كما نلاحظ أن قيم معامل الارتباط فوق القطر مساوية تماماً للقيم تحت القطر.

يشير الجدول السابق إلى وجود ارتباط ذي دلالة إحصائية بين الموقف من المدينة ومدة الإقامة فيها

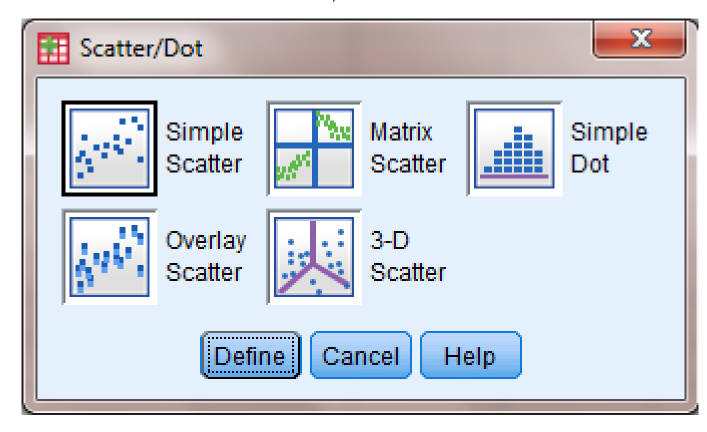
sig< 0.001 (أي أن بإمكاننا رفض فرضية العدم وقبول الفرضية البديلة). ويشير معامل بيرسون *r =* *0.935* إلى وجود علاقة ارتباط قوية جداً وموجبة بين هذين المتغيرين، أي أن الموقف من المدينة ومدة الإقامة في المدينة يسيران معاً فيزداد الموقف من المدينة مع ازدياد مدة الإقامة في المدينة.

ويمكن كما ذكرنا تمثيل أو اختبار احتمال وجود العلاقة الخطية بين المتغيرين من خلال استخدام منحنى الانتشار *Scatter Plot* وذلك كمايلي:

Graphs←Legacy Dialog←Scatter/Dot...

الشكل 6.7. كيفية إظهار منحنى الانتشار

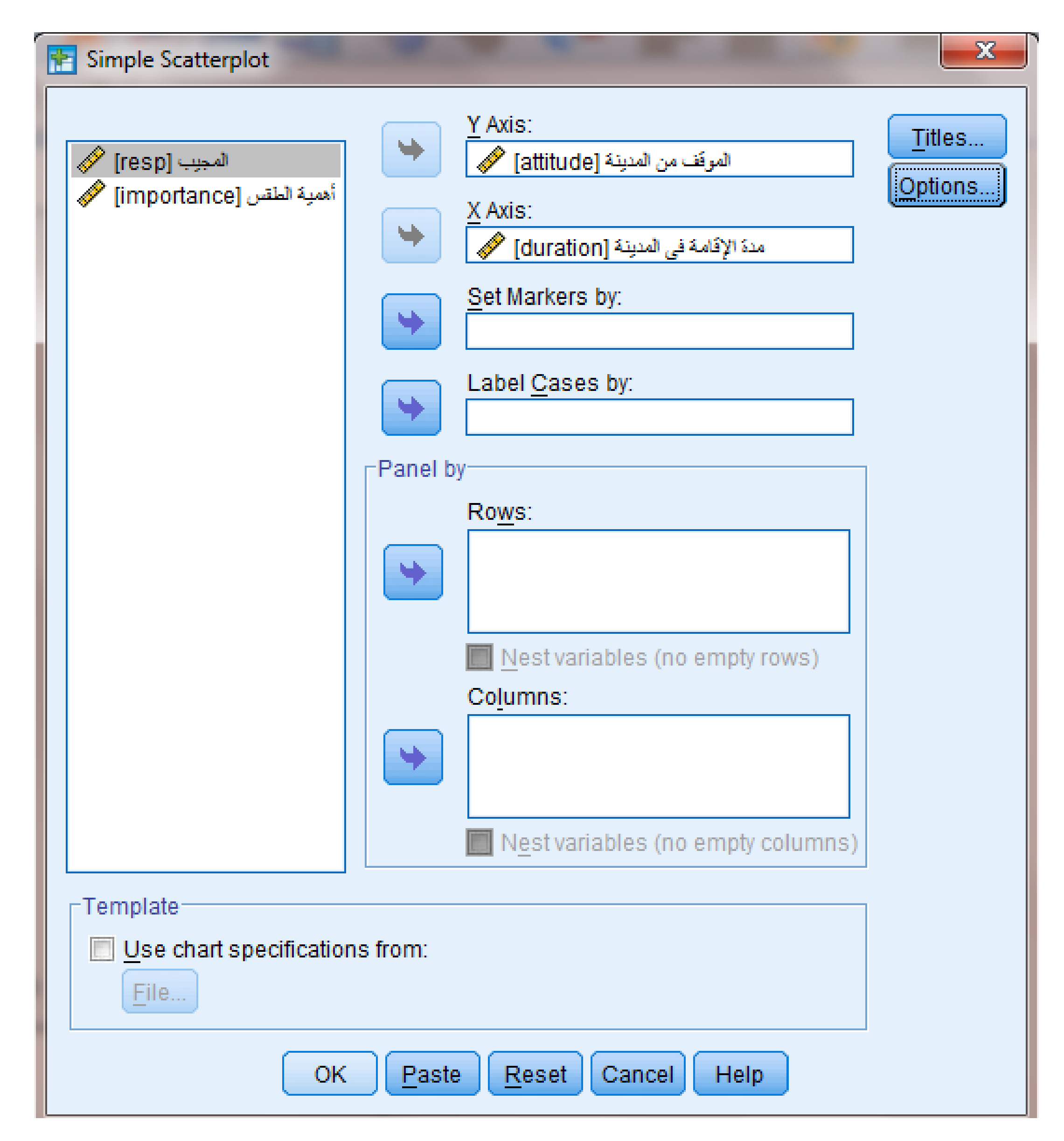
في النافذة *Scatter/Dot* نختار *Simple Scatter* ثم ننقر فوق *Define*.



الشكل 7.7. النافذة *Simple scatter*

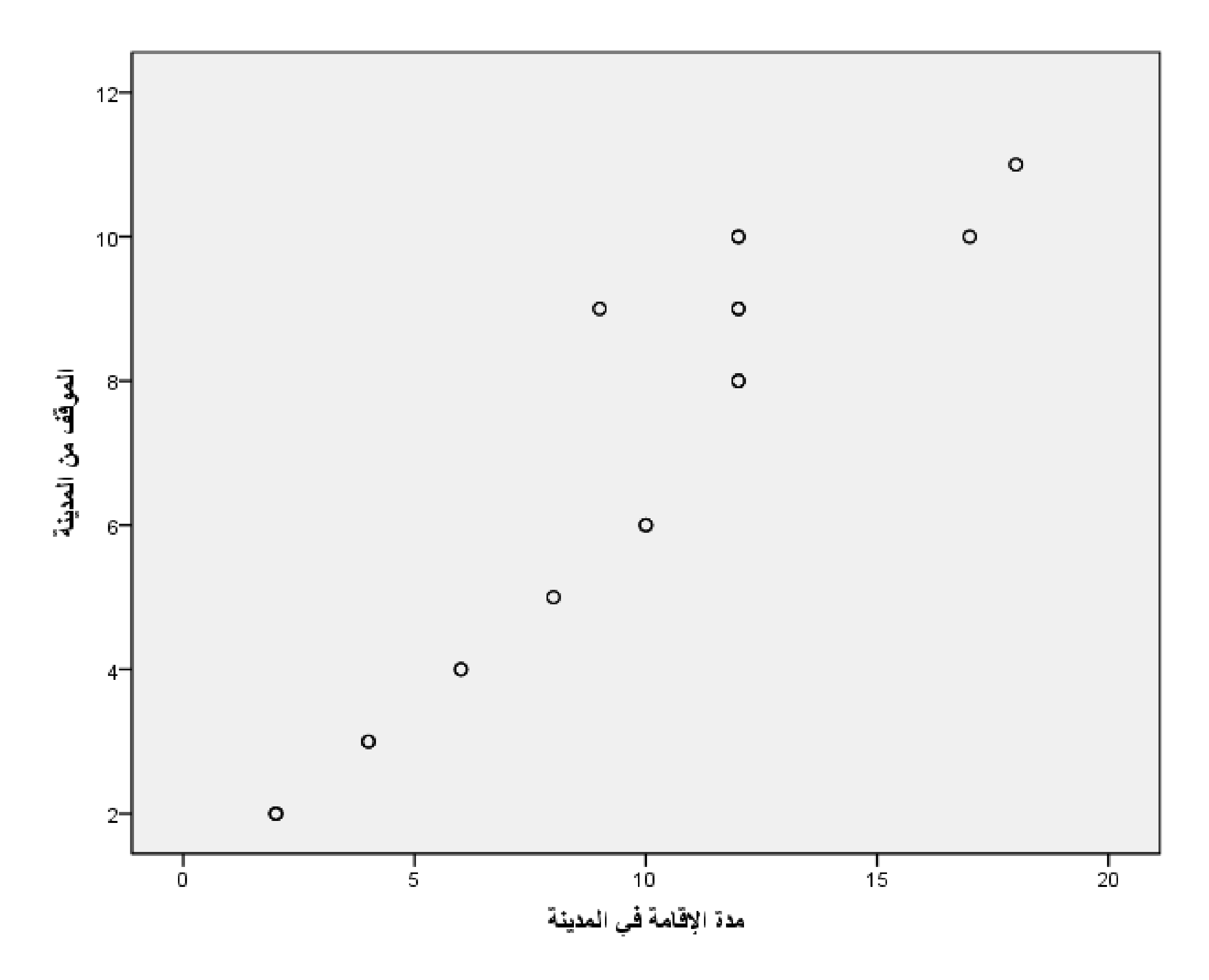
في النافذة *Simple Scatterplot* نضع الموقف من المدينة في المربع *Y Axis* والمتغير مدة الإقامة

في المربع *X Axis*.



الشكل 8.7. النافذة *Simple Scatterplot*

بالنقر فوق *OK* ينتج لدينا الشكل البياني التالي الذي يبين توزع إحداثيات العلاقة بين الموقف من المدينة ومدة الإقامة في المدينة.



الشكل 9.7. منحنى الانتشار للعلاقة بين الموقف من المدينة ومدة الإقامة في المدينة

يظهر الشكل السابق وجود علاقة خطية واضحة بين المتغيرين المدروسين (الموقف من المدينة ومدة الإقامة) حيث تتجمع النقاط في الشكل على خط مستقيم ذي ميل إيجابي. ويظهر من توضع النقاط كيف أن تزايد الموقف من المدينة يترافق مع تزايد مدة الإقامة في المدينة.

في النهاية، نشير إلى أن مصفوفة الارتباط الخطي يمكن أن تتضمن أكثر من متغير وفي هذه الحالة نقرأ النتائج في المثلث الأدنى أو الأعلى من الجدول (فوق قطر الجدول أو تحته) حيث أن النتائج في الجزء الأدنى من الجدول (المثلث الأدنى) هي انعكاس أو هي ذاتها في الجزء الأعلى من الجدول. وبما أن معامل الارتباط على القطر يساوي دائماً (1) يميل معظم الكتاب إلى عدم إظهار قيم معامل الارتباط على القطر.

	V_1	V_2	V_3	V_4	V_5
V_1					
V_2	0.5				
V_3	0.3	0.4			
V_4	0.1	0.3	0.6		
V_5	0.2	0.5	0.3	0.7	

الشكل 10.7. كيفية عرض مصفوفة الارتباط في الكتابة العلمية

قائمة المراجع

المراجع العربية:

- الساعاتي، عبدالرحيم؛ حسن، أحمد السيد؛ حابس، عصام؛ البحطيطي، عبدالرحيم؛ أبو العلا، لبنى؛ الشربيني، زكريا (2009)، **تطبيقات في التحليل الإحصائي للعلوم الإدارية والإنسانية،** الطبعة الثانية، جامعة الملك عبدالعزيز، جدة.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS*،** الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Buglear J. (2003),* ***Stats means Business: A Guide to Business Statistics****, 3rd Edition, Elsevier Butterworth-Heinemann, England.*
- *Coakes S.J. (2005),* ***SPSS for Windows: Analysis without Anguish****, John Wiley, Australia.*
- *Field A. (2006),* ***Discovering Statistics Using SPSS****, 2nd Edition, SAGE, England.*
- *Hair J.F. JR., Bush R.P., &Ortinau D.J. (2003),* ***Marketing Research Within a Changing Information Environment****, McGraw-Hill/Irwin, 2nd Edition, USA.*
- *Ho R. (2006),* ***Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS****, Chapman & Hall/CRC, USA.*
- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*
- *Malhotra N.K. &Briks D.F. (2007),* ***Marketing Research: An Applied Approach****, 3rd European Edition, Pearson Education Limited, Italy.*
- *Morgan G.A., Leech N.L., Gloeckner G.W., & Barrett K.C. (2004),* ***SPSS for Introductory Statistics: Use and Interpretation****, 2nd Edition, Lawrence Erlbaum Associates, USA.*
- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows****, 3rd Edition, McGraw Hill, USA.*
- ***SPSS Statistics Base 17.0 User's Guide (2007)****, SPSS Inc., USA.*

مقترحات وتمارين للفصل السابع

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1- كيف يمكن وصف العلاقة بين متغيرين؟

(الحل في الفقرة:1.)

2- متى يستخدم معامل ارتباط سبيرمان؟

(الحل في الفقرة:2.)

3- ما هي شروط اختبار الارتباط الخطي؟

(الحل في الفقرة:1.2.)

4- ما هي دلالة معامل بيرسون؟

(الحل في الفقرة:2.2.)

5- قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.3. (الفصل الثالث) ثم قم باختبار إمكانية وجود علاقة خطية بين الموقف من الانترنت والموقف من التكنولوجيا.

الفصل الثامن تحليل الانحدار الخطي

Chapter 8: Linear regression analysis

الكلمات المفتاحية:

الانحدار الخطي البسيط، الانحدار الخطي المتعدد، معامل التحديد.

ملخص:

يشرح الفصل كيفية استخدام تحليل الانحدار الخطي للتنبؤ بتغيرات متغير تابع بدلالة متغير أو مجموعة متغيرات مستقلة. ويميز الفرق بين تحليل الانحدار الخطي البسيط الذي يختبر أثر متغير مستقل واحد في متغير تابع واحد والانحدار الخطي المتعدد الذي يختبر أثر أكثر من متغير مستقل في متغير تابع واحد.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- إدراك كيفية اختبار أثر متغير كمي واحد أو أكثر في متغير كمي تابع.
- الإلمام بكيفية تطبيق تحليل الخطي البسيط والمتعدد في *SPSS*.
- فهم الفرق بين الانحدار الخطي البسيط والمتعدد.
- فهم كيفية تفسير مخرجات تحليل الانحدار الخطي البسيط والمتعدد.

الفصل الثامن: تحليل الانحدار الخطي

يعتبر تحليل الانحدار أحد الأساليب الإحصائية المهمة والتي تستخدم بشكل واسع منذ القرن التاسع عشر لتحديد وتوضيح التأثيرات بين المتغيرات المستقلة *independent variables* والمتغير التابع *dependent variable*. يستخدم الانحدار أيضاً للتنبؤ بقيمة المتغير التابع بدلالة المتغيرات المستقلة بعد إيجاد معادلة الانحدار.

وكما سبق وشرحنا في الفصل السابق فإن نموذج العلاقة بين المتغيرات قد يكون خطياً أو غير خطي إلا أننا في هذا المقرر سنركز على النماذج الأكثر انتشاراً أي النماذج الخطية. فإذا كان هناك متغير مستقل واحد ومتغير تابع واحد فيسمى التحليل بالانحدار الخطي البسيط *Simple Linear Regression*. أما إذا كان هناك أكثر من متغير مستقل ومتغير تابع واحد فيسمى التحليل بالانحدار الخطي المتعدد *Multiple Linear Regression*.

1- الانحدار الخطي البسيط *Simple linear regression*

كما رأينا في الفصل السابق فإن تحليل الارتباط الخطي يفيد في قياس قوة العلاقة الخطية بين متغيرين إلا أنه لا يعطينا معادلة هذه العلاقة. فإذا ما أردنا معرفة أثر المتغير المستقل في المتغير التابع فإن علينا اللجوء إلى الانحدار الخطي البسيط *Simple Linear regression*.

وكثيراً ما يتم إطلاق اسم العلاقة الدالية أو الدوال على العلاقة السببية بين المتغير المستقل والتابع. ومن أمثلة هذه الدوال نذكر:

- دالة الطلب: تفترض أن الكمية المطلوبة من سلعة (المتغير التابع) تتأثر بالتغيرات في سعر السلعة (المتغير المستقل).
- دالة العرض: الكمية المعروضة من السلعة (المتغير التابع) تتأثر بالتغيرات في سعر السلعة (المتغير المستقل).
- دالة الاستهلاك: الإنفاق الاستهلاكي (المتغير التابع) يتأثر بالتغيرات في دخل الأفراد (المتغير المستقل).

يمكننا الانحدار الخطي البسيط إذاً من إيجاد معادلة المستقيم الأكثر تمثيلاً للعلاقة السببية بين قيم

المتغيرين (المستقل والتابع) أي أنه يسمح لنا بالتوصل إلى الأشكال الرياضية المحددة للعلاقات السببية أو الدوال.

1-1. معادلة الانحدار الخطي البسيط *Simple linearregression formula*

يأخذ نموذج الانحدار الخطي البسيط الصيغة التالية:

$$Y = \beta_0 + \beta_1 X$$

حيث:

- Y : المتغير التابع
- X : المتغير المستقل
- $\beta 0$: الحد الثابت أو معلمة تقاطع خط الانحدار (المستقيم) مع المحور العمودي (محور العينات).
- β_1**:** يمثل معلمة الميل. تكون إشارة معلمة الميل موجبة (+) إذا كانت العلاقة طردية بين المتغير المستقل والتابع أي أن زيادة قيمة المتغير المستقل تؤدي إلى ارتفاع قيمة المتغير التابع، أما إذا كانت العلاقة عكسية بين المتغيرين فتكون إشارة معلمة الميل سالبة (−).

نعني إذاً بالمعادلة الخطية أن المتغير التابع Y هو دالة خطية من المتغير المستقل X، أي أن زيادة وحدة واحدة إلى X سوف تؤدي إلى التغير في كمية Y. ويتم تحديد التغير في Y بدلالة X من خلال تحديد الثوابت $\beta 0$ و $\beta 1$ والتي تدعى بالمعلمات أو بمعاملات الانحدار *Coefficients*.

2-1. شروط الانحدار الخطي البسيط *Assumptions of simple linearregression*

من الشروط الأساسية الواجب توافرها لتطبيق الانحدار الخطي البسيط:

- وجود علاقة خطية بين المتغيرين المستقل X والتابع Y. وقد سبق وشرحنا في الفصل السابق كيفية استكشاف إمكانية وجود العلاقة الخطية من خلال منحنى الانتشار *Scatter Plot*.
- أن يكون توزيع المتغير التابع والمتغير المستقل توزيعاً طبيعياً *Normal Distribution*.
- وبشكل عام يتوجب أن تكون بيانات المتغيرين المستقل والتابع من النمط المدرج *Interval* أو النسب *Ratio*. ويمكن في بعض الأحيان استخدام الانحدار مع مستوى قياس أدنى للمتغير المستقل.

3-1. الاختبار باستخدام SPSS: Test using SPSS

في مثال الموقف من المدينة الذي تم شرحه في الفصل السابق أراد الباحث اختبار أثر مدة الإقامة في المدينة (متغير مستقل) في الموقف من المدينة (متغير تابع). أي أن الباحث يسعى إلى إيجاد معادلة الانحدار الخطي التالية (إن وجدت طبعاً):

الموقف من المدينة = $ß_0$ + $ß_1$*مدة الإقامة في المدينة

يمكن هنا صياغة فرضية العدم والفرضية البديلة كمايلي:

فرضية العدم: لا تؤثر مدة الإقامة في المدينة في الموقف من المدينة، أي أن:

$$H_0: ß_1 = 0$$

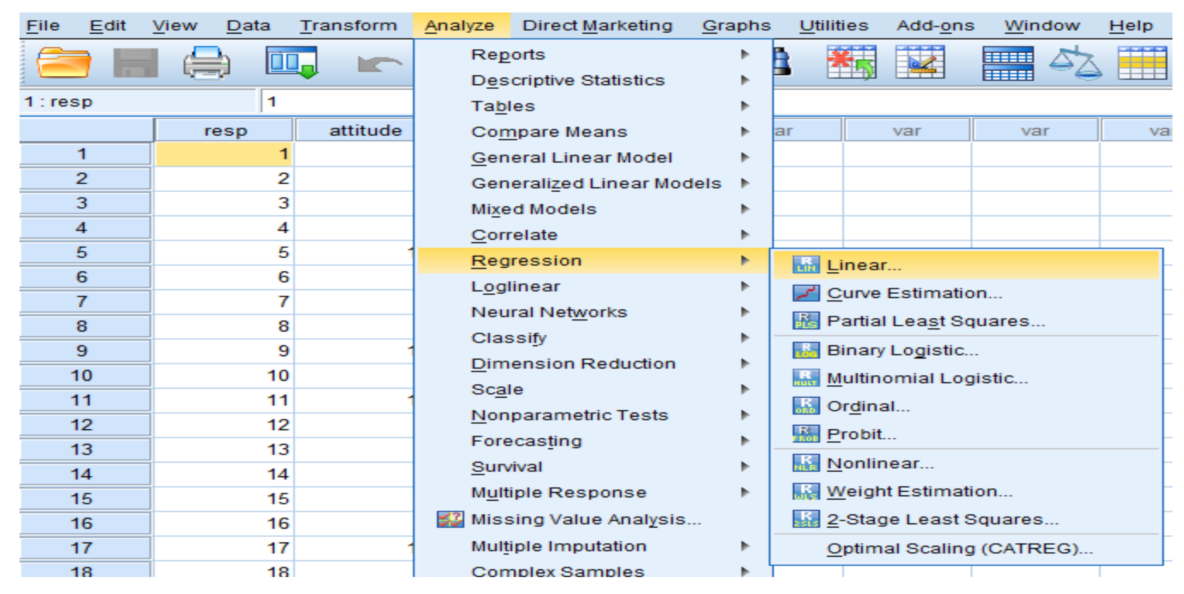
الفرضية البديلة: تؤثر مدة الإقامة في المدينة في الموقف من المدينة، أي أن:

$$H_1: ß_1 \neq 0$$

ويستخدم اختبار t لاختبار معلمة الميل، حيث تستخدم قيمة *p-value* (أو *sig*) المرافقة لإحصائية t للمعلمة لاختبار إمكانية رفض فرضية العدم. فإذا كانت $p\text{-value} < 0.05$ نرفض فرضية العدم عند مستوى دلالة 5% ونقبل الفرضية البديلة، أي أن المتغير المستقل يؤثر في المتغير التابع. أما إذا كانت $p\text{-value} > 0.05$ فلا يمكننا رفض فرضية العدم وبالتالي لا يوجد أثر ذو دلالة إحصائية للمتغير المستقل في المتغير التابع.

ولتطبيق اختبار الانحدار الخطي البسيط في *SPSS* نتبع المسار التالي:

Analyze←Regression←Linear



الشكل 1.8. كيفية تطبيق تحليل الانحدار الخطي في *SPSS*

في النافذة *Linear Regression* نضع المتغير المستقل "مدة الإقامة في المدينة" ضمن المربع *Independent(s)* ونضع المتغير التابع "الموقف من المدينة" ضمن المربع *Dependent*.

ونلاحظ في القائمة المنسدلة *Method* أن هناك العديد من طرق الانحدار الخطي، مثل:

- *Enter*: الطريقة الافتراضية ضمن *SPSS* حيث يتم إدخال كافة المتغيرات المستقلة في النموذج (معادلة الانحدار) ليقوم الباحث باختبار أثرها مجتمعة في المتغير التابع.
- *Stepwise*: يتم بموجبها إدخال أولاً المتغير الذي يتصف بأعلى معنوية (دلالة إحصائية) في علاقته مع المتغير التابع يليه المتغير الثاني من ناحية المعنوية وهكذا. تتيح هذه الطريقة متابعة التغيرات التي تطرأ على النموذج عند إضافة كل متغير ذي أثر معنوي.
- *Backward*: تتناول هذه الطريقة جميع المتغيرات ومن ثم تبدأ باستبعاد المتغيرات مبتدأة من الأكثر "لا معنوية" ومن ثم الذي يليه من حيث عدم المعنوية وهكذا لغاية التوقف عند المتغيرات التي تستوفي درجة المعنوية (الدلالة الإحصائية).
- *Forward*: تقوم هذه الطريقة بإدخال كافة المتغيرات بادئة بالمتغير الأكثر معنوية أولاً ومن ثم الذي يليه من حيث المعنوية ليتم التوقف عند عدم استيفاء المتغير لدرجة المعنوية (الدلالة الإحصائية) المقررة.

ISSN: 2617-989X

اختبار ANOVA يشير إلى إمكانية استخدام نموذج الانحدار (من خلال المتغير المستقل الذي تم



الشكل 2.8. النافذة *Linear Regression*

بالنقر فوق *OK* تظهر نتائج تحليل الانحدار حيث يشير الجدول الأول إلى الطريقة المعتمدة في التحليل.

جدول 1.8. المتغيرات الداخلة في التحليل

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	مدة الإقامة في المدينة[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: الموقف من المدينة

ويظهر جدول *ANOVA* قيمة معامل *F* لنموذج الانحدار الذي يبين الدلالة الإحصائية لنموذج الانحدار الكلي. نلاحظ في هذا الجدول أن $F(1,28)=194.188$ و $sig < 0.001$ وبالتالي أصغر من α أي أن اختبار *ANOVA* يشير إلى إمكانية استخدام نموذج الانحدار (من خلال المتغير المستقل الذي تم إدخاله) لشرح تغيرات المتغير التابع.

جدول 2.8. اختبار *ANOVA* في تحليل الانحدار الخطي البسيط

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	245.779	1	245.779	194.288	.000[a]
	Residual	35.421	28	1.265		
	Total	281.200	29			

a. Predictors: (Constant), مدة الإقامة في المدينة

b. Dependent Variable: الموقف من المدينة

ويمكن تحديد نموذج أو معادلة الانحدار من خلال الجدول *Coefficients* الذي يظهر معاملات الانحدار.

جدول 3.8. معاملات الانحدار *Coefficients*

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.950	.454		2.090	.046
	مدة الإقامة في المدينة	.605	.043	.935	13.939	.000

a. Dependent Variable: الموقف من المدينة

يظهر الجدول *Coefficients* أن قيمة *p-value*لمعلمة الميل أي لمعامل الانحدار الخطي للمتغير المستقل أصغر من 0.001 (أي أصغر من α) وبالتالي يمكننا رفض فرضية العدم وقبول الفرضية البديلة أي أن مدة الإقامة في المدينة تؤثر بشكل ذي دلالة إحصائية في الموقف من المدينة.

يمكننا الجدول *Coefficient*من استنتاج نموذج الانحدار الخطي (كتابة معادلة الانحدار الخطي) التالي:

الموقف من المدينة = 0.950 + 0.605*مدة الإقامة في المدينة

تشير معلمة الميل *B1* إلى أن زيادة مدة الإقامة لسنة واحدة يؤدي إلى زيادة الموقف من المدينة بمقدار 0.605.

أما معلمة *Standardized Coefficient* والتي يشار إليها بالجدول أعلاه بـ *Beta*، فهي معلمة الميل للنموذج المقدر باستعمال القيم المعيارية لكل من المتغير المستقل والتابع بدل القيم الأصليةحيث تحسب

هذه القيم المعيارية وفق المعادلة التالية:

$$(X - \bar{X})/S$$

ولا يحتوي هذا النموذج على معلمة التقاطع (الحد الثابت) *B0*.

جدول 4.8. معامل التحديد

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.935[a]	.874	.870	1.125

a. Predictors: (Constant), مدة الإقامة في المدينة

وأخيراً يتضمن الجدول *Model Summary* أحد أهم المؤشرات لنموذج الانحدار وهو **معامل التحديد** ***Coefficient of Determination*** ويرمز له R^2 والذي يعتبر مقياساً لجودة تفسير النموذج. يدل هذا المعامل على النسبة المئوية لتغيرات المتغير التابع المشروحة بدلالة نموذج الانحدار (أي بدلالة المتغير المستقل في حالة الانحدار الخطي البسيط نظراً لوجود متغير مستقل واحد في النموذج). تتراوح قيمة معامل التحديد بين 0 و 1 ويمكن القول إنه كلما اقتربت قيمة R^2 من 1 كلما دلّ ذلك على جودة النموذج.

وفي نموذج الانحدار الخطي البسيط هناك علاقة بين معامل التحديد ومعامل الارتباط الخطي لبيرسون r بين المتغير المستقل والتابع حيث:

$$r = \sqrt{R^2}$$

يتصف معامل التحديد بأنه لو أضيف متغير مستقل للنموذج فإن قيمته سترتفع حتى لو لم تكن هناك أهمية للمتغير المستقل في النموذج. ولهذا يتم احتساب معامل التحديد المصحح *Adjusted* وقيمته دائماً أقل من قيمة معامل التحديد (غير المصحح).

أما الخطأ المعياري للتقدير *Standard Error of Estimate* فيقيس تشتت القيم المشاهدة عن خط الانحدار، ويعني الحصول على قيمة صغيرة لهذا المؤشر صغر الأخطاء العشوائية وبالتالي جودة تمثيل خط الانحدار لنقاط شكل الانتشار (ونادراً ما يشار إلى هذه المؤشر عند شرح النتائج).

وبالعودة إلى نتيجة التحليل نلاحظ أن قيمة R^2 تساوي 0.874 مما يعني أن نموذج الانحدار الخطي (أو أن المتغير المستقل – مدة الإقامة في المدينة) يشرح 87.4% من التغيرات في قيم المتغير التابع

(الموقف من المدينة). وترجع 12.6% (المتبقية) من التغيرات غير المفسرة إلى عوامل عشوائية كأن تكون هناك متغيرات مهمة لم تضمن في النموذج.

2- الانحدار الخطي المتعدد *Multiple linear regression*

يعتبر الانحدار الخطي المتعدد *Multiple Linear Regression* امتداداً للانحدار الخطي البسيط إلا أنه يأخذ في الحسبان أكثر من متغير مستقل. ويستخدم الانحدار الخطي المتعدد للتنبؤ بقيمة المتغير التابع بدلالة مجموعة من المتغيرات المستقلة. ويتم ذلك من خلال تمثيل العلاقة بين المتغير التابع Y والمتغيرات المستقلة على شكل معادلة خطية تكتب بالصورة التالية:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

حيث $B_1, B_2, \dots, B_n$ تمثل معاملات الانحدار الجزئية *Partial Regression Coefficient* أو الميول الجزئية كما يدعوهاالبعض معاملات المتغيرات المستقلة.

1.2. شروط الانحدار الخطي المتعدد *Assumptionsof multiple linear regression*

إن شروط النموذج الخطي المتعدد هي نفسها شروط النموذج الخطي البسيط يضاف إليها وجوب ألا تكون المتغيرات المستقلة مرتبطة مع بعضها البعض بشكل كبير، لأن ذلك يولد مشكلة الارتباط الخطي المتعدد *Multicollinearity*. ويمكن التحقق من هذا الشرط من خلال حساب معاملات الارتباط *pearson correlation* بين المتغيرات المستقلة والتأكد من أنها ليست مرتبطة بشكل كبير فيما بينها (لا تتجاوز أو لا تساوي 0.90).

2.2. الاختبار باستخدام SPSS: *Test using SPSS*

في مثالنا السابق أضاف الباحث المتغير المستقل "أهمية الطقس" إلى نموذج الانحدار الخطي وأعاد إجراء الاختبار وفق الخطوات المبينة سابقاً فحصل على النتائج التالية:

جدول 5.8. نتيجة الاختبار الخطي المتعدد

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	أهمية الطقس, مدة الإقامة في المدينة[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: الموقف من المدينة

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.974[a]	.948	.944	.738

a. Predictors: (Constant), أهمية الطقس, مدة الإقامة في المدينة

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	266.510	2	133.255	244.929	.000[a]
	Residual	14.690	27	.544		
	Total	281.200	29			

a. Predictors: (Constant), أهمية الطقس, مدة الإقامة في المدينة

b. Dependent Variable: الموقف من المدينة

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		Double-click to activate	Error	Beta		
1	(Constant)	[illegible]	.309		1.432	.164
	مدة الإقامة في المدينة	.485	.035	.749	14.056	.000
	أهمية الطقس	.276	.045	.329	6.173	.000

a. Dependent Variable: الموقف من المدينة

يشير جدول *ANOVA* إلى إمكانية استخدام نموذج الانحدار الخطي المتعدد لشرح تغيرات المتغير التابع "الموقف من المدينة" حيث أن $F(2,27)= 244.929$ و $sig<0.001$.

ومن خلال النتائج الظاهرة في الجدول *coefficients* يمكن كتابة نموذج الانحدار الخطي كمايلي:

الموقف من المدينة = 0.443 + 0.485*مدة الإقامة في المدينة + 0.276*أهمية الطقس

تشير معلمة مدة الإقامة في المدينة إلى أن زيادة مدة الإقامة لسنة واحدة سيؤدي إلى تحسين الموقف من

المدينة بمقدار 0.485 بافتراض ثبات المتغير "أهمية الطقس". فيما ستؤدي زيادة أهمية الطقس بمقدار درجة واحدة مع ثبات مدة الإقامة في المدينة إلى زيادة الموقف من المدينة بمقدار 0.276. وبالتالي فإن نموذج الانحدار الخطي المتعدد يشير إلى أن المتغير "مدة الإقامة في المدينة" يؤثر في "الموقف من المدينة" بشكل أكبر من المتغير "أهمية الطقس".

كما يشير معامل التحديد R^2 إلى أن نموذج الانحدار يشرح 94% من تغيرات المتغير التابع (أي أن المتغيرين المستقلين معاً يشرحان 94% من تغيرات المتغير التابع "مدة الإقامة في المدينة").

قائمة المراجع

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، **أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام *SPSS***، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- الساعاتي، عبدالرحيم؛ حسن، أحمد السيد؛ حابس، عصام؛ البحطيطي، عبدالرحيم؛ أبو العلا، لبنى؛ الشربيني، زكريا (2009)، **تطبيقات في التحليل الإحصائي للعلوم الإدارية والإنسانية**، الطبعة الثانية، جامعة الملك عبدالعزيز، جدة.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة *SPSS***، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- *Coakes S.J. (2005),* ***SPSS for Windows: Analysis without Anguish****, John Wiley, Australia.*
- *Hair J.F. JR., Bush R.P., &Ortinau D.J. (2003),* ***Marketing Research Within a Changing Information Environment****, McGraw-Hill/Irwin, 2nd Edition, USA.*
- *Ho R. (2006),* ***Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS****, Chapman & Hall/CRC, USA.*
- *Landau S. &Everitt B.S. (2004),* ***A Handbook of Statistical Analysis Using SPSS****, Chapman & Hall/CRC Press, USA.*
- *Leech N.L., Barrett K.C., & Morgan G.A. (2005),* ***SPSS for Intermediate Statistics: Use and Interpretation****, 2nd Edition, Laerence Erlbaum Associates, Inc., USA.*
- *Malhotra N.K. (2010),* ***Marketing Research: An Applied Orientation****, 6th Edition, Pearson, USA.*
- *Malhotra N.K. &Briks D.F. (2007),* ***Marketing Research: An Applied Approach****, 3rd European Edition, Pearson Education Limited, Italy.*
- *Pallant J. (2007),* ***SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows****, 3rd Edition, McGraw Hill, USA.*
- ***SPSS Statistics Base 17.0 User's Guide (2007)****, SPSS Inc., USA.*

مقترحات وتمارين للفصل الثامن

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1. **ما الفرق بين الانحدار الخطي البسيط والانحدار الخطي المتعدد؟**

   (الحل في الفقرة: مقدمة الفصل)

2. **ما هي معادلة الانحدار الخطي البسيط؟**

   (الحل في الفقرة:1.1.)

3. **ما هي شروط اختبار الانحدار الخطي البسيط؟**

   (الحل في الفقرة:2.1.)

4. **ما هو معامل التحديد؟**

   (الحل في الفقرة:3.1.)

5. **ما هي شروط اختبار الانحدار الخطي المتعدد؟**

   (الحل في الفقرة:1.2.)

6. **قم بإنشاء ملف *SPSS* جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 2.7.(الفصل السابع) ثم قم باختبار مايلي:**
   - **أثر أهمية الطقس في الموقف من المدينة.**
   - **أثر أهمية الطقس ومدة الإقامة في المدينة في الموقف من المدينة.**